The approximation of the normalized empirical distribution function by a Brownian bridge

by Péter Major
Mathematical Institute of the Hungarian Academy of Sciences

Abstract: In this work an (asymptotically) optimal approximation of the normalized empirical distribution function by a Brownian bridge is presented in the form of a series of problems. The formulation of the problems and their solutions are separated in this text. The discussion is based on the paper *An approximation of partial sums of independent RV's and the sample DF. I.* written by János Komlós, Péter Major and Gábor Tusnády which appeared in the journal *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **32** (1975), pages 111–131. The main difference between this text and the original paper is that here the details are worked out thoroughly, all "it is easy to see that ... " type arguments are omitted. The problems formulated in this work are of different level of difficulty. My main goal was to give a complete and understandable explanation. I discussed some technical problems in detail and tried to explain the ideas behind them even if this caused sometimes certain repetition and made the text longer.

The formulation of the main problem

In this series of problems we investigate the approximation of the normalized empirical distribution function by a Brownian bridge. Let us formulate this problem in more details.

Let $B(t) = B(t,\omega)$, $0 \le t \le 1$, be a Brownian bridge on a probability space $(\Omega, \mathcal{B}, P)$, i.e. let $B(t)$ be a Gaussian process with expectation $EB(t) = 0$ and covariance function $EB(s)B(t) = \min(s,t) - st$, $0 \le s,t \le 1$. Furthermore, we assume that the trajectories $B(\cdot,\omega)$ of the Brownian bridge are continuous functions on the interval $[0,1]$ for all elementary events $\omega \in \Omega$. It follows from classical results of the probability theory that Brownian bridges with the above properties really exist.

Let $P_n(t)$, $0 \le t \le 1$, denote the empirical distribution function corresponding to a sample consisting of n independent and on the interval $[0,1]$ uniformly distributed random variables, i.e. let $\zeta_1, \dots, \zeta_n$ be a sequence of independent and on the interval $[0,1]$ uniformly distributed random variables, and put $P_n(t) = \dfrac{1}{n} \sum\limits_{j=1}^{n} I(\{\zeta_j \le t\})$, $0 \le t \le 1$, where $I(A)$ denotes the indicator function of the set A. Let

$$Z_n(t) = \sqrt{n}[P_n(t) - t] \ , \quad 0 \le t \le 1,$$

be its standardization which we shall further call the standardized empirical distribution function. The covariance function of the processes $Z_n(t)$ and $B(t)$, $0 \le t \le 1$, agree. We want to show that these two processes can be put close to each other with the help of an appropriate construction. More explicitly, we want to prove the following result:

Approximation Theorem. *Let a Brownian bridge $B(t)$ be given on a sufficiently rich probability space $(\Omega,\mathcal{B},P)$. Then for all numbers $n=2,3,\dots$ a sequence $\zeta_1,\dots,\zeta_n$ of independent and on the interval $[0,1]$ uniformly distributed random variables can be constructed in such a way that the empirical distribution function $P_n(t)=\frac{1}{n}\sum\limits_{k=1}^n I(\{\zeta_n\le t\})$ and its standardization $Z_n(t)=\sqrt{n}[P_n(t)-t]$, $0\le t\le 1$, made with the help of these random variables $\zeta_1,\dots,\zeta_n$ satisfy the relation*

$$P\left(\sqrt{n}\sup_{0\le t\le 1}|Z_n(t)-B(t)|>C_1\log n+x\right)\le C_2e^{-\lambda x}$$

for all numbers $x>0$ with some universal (independent of the parameter n) constants $C_1>0$, $C_2>0$ and $\lambda>0$.

Remark: The probability space where we want to make a construction satisfying the *Approximation Theorem* is sufficiently rich if there exists a sequence η_k, $k=1,2,\dots$, of independent and on the interval $[0,1]$ uniformly distributed random variables on it which is also independent of the Brownian bridge $B(t)$. At the expense of some extra-work such a modified construction could be made which applies no extra random variables beside the Browian bridge $B(t)$. But it is more convenient to carry out the construction presented in his work, and the extra-condition we have imposed is not a serious restriction in possible applications of the result.

In the construction satisfying the above *Approximation Theorem* we shall construct the empirical distribution function $P_n(t)$ as an appropriate transform of the Brownian bridge in such a way that its normalization $Z_n(t)$ is close to the Brownian bridge $B(t)$. The method of this construction is an appropriate adaptation of the quantile transform to the present problem. The main difficulty is caused by the fact that the multi-dimensional distributions of an empirical distribution function are also prescribed, while the quantile transform only deals with the construction of random variables with a prescribed one-dimensional distribution. To overcome this difficulty we shall construct the (standardized) empirical distribution function $Z_n(t)$ by means of a fixed Brownian bridge $B(t)$ subsequently for newer and newer points t in such a way that the random variable $Z_n(t)$ has that conditional distribution which the values of the previously constructed random variables $Z_n(s)$ prescribe. We do this subsequently, and in the l-th step we define the process $Z_n(t)$ in the diadic points $t=k2^{-l}$, $k=0,1,\dots,2^l$. In the present problem the conditional distributions we have to work with can be well handled. The construction gives a good approximation because in the l-th step we define the values of the process in all points $t_{k,l}=k2^{-l}$, $1\le k\le 2^l$, which is a relatively dense subset of the interval $[0,1]$. If the supremum of the process $\sqrt{n}(Z_n(t)-B(t))$ in the already constructed points increases in each step relatively little, then this method provides a good approximation. By working out the details we can show that roughly speaking the above supremum increases only with a constant in each step. In such a way we get a construction for which the supremum of the difference $\sqrt{n}(Z_n(t)-B(t))$ is less than const. $\log n$ with probability almost one. It is worth mentioning that this result is sharp. The difference $\sqrt{n}(Z_n(t)-B(t))$ is greater than const. $\log n$ with (another) appropriate positive constant for an arbitrary construction. The proof of this statement is considerably simpler than the proof of the *Approximation Theorem*, and we shall discuss it in another series of problems.

To apply the quantile transform method in our problem we need a sharp limit theorem for the conditional distribution function of the random variables $Z_n(t_2)-Z_n(t_1)$ under appropriate conditions. Moreover, to get sharp results it is not sufficient to have a good limit theorem in the usual way. We need such large deviation type results which also describe the goodness of approximation in the case of non-typical values or non-typical conditions. We formulate a classical result of the large deviation theory in the special case we need it. Then we formulate two problems whose solutions supply a good bound on the goodness of the approximation of the quantile transformation in the case we need it.

We shall apply the following notations. Let $\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ denote the standard normal density and $\Phi(x)=\int_{-\infty}^{x}\varphi(x)\,dx$ the standard normal distribution function. We shall need the following result.

Theorem. *Let $F_n(x)$ be the standardization of the binomial distribution $B\left(n,\frac12\right)$ with parameters n and $p=\frac12$, i.e. let $F_n(x)=P\left(2\dfrac{\eta_1+\cdots+\eta_n-\frac n2}{\sqrt{n}}<x\right)$, where $\eta_1,\dots,\eta_n$ are independent and identically distributed random variables, $P(\eta_j=1)=P(\eta_j=0)=\frac12$, $j=1,\dots,n$. Then there exist universal (independent of n) constants $K>0$ and $A>0$ in such a way that the distribution function $F_n(x)$ satisfies the following inequalities in the interval $|x|\le A\sqrt{n}$:*

$$(1-\Phi(x))\exp\left\{-\frac{K(x^3+1)}{\sqrt{n}}\right\}\le 1-F_n(x)\le(1-\Phi(x))\exp\left\{\frac{K(x^3+1)}{\sqrt{n}}\right\}$$
$$\Phi(-x)\exp\left\{-\frac{K(x^3+1)}{\sqrt{n}}\right\}\le F_n(-x)\le\Phi(-x)\exp\left\{\frac{K(x^3+1)}{\sqrt{n}}\right\},\qquad \text{if } 0\le x\le A\sqrt{n}.$$

(The statement of this theorem actually holds for all constants $A<\frac12$ and appropriate $K=K(A)$, but we shall not need this sharper result.)

Let us remark that the above theorem is the special case of a more general result. A $B(n,\frac12)$ distributed random variable, as we have remarked in the formulation of this theorem, can be represented as the sum of n independent $B(1,\frac12)$ distributed random variables. It follows from the general theory of the large deviations that the estimation formulated in the theorem also holds for normalized sums of independent identically distributed random variables if the summands have moment generating function, i.e. if the distribution function $F(x)$ of the summands satisfies the condition $\int e^{tx}F(\,dx)<\infty$ if $|t|<a$ with some appropriate constant $a>0$. This result and its proof can also be found in problem 22 of the series of problem *The Theory of Large Deviation I.* in my homepage. (At present it exists only in Hungarian.)

1. Problems

1.) There exist such constants $C_1 > 0$ and $C_2 > 0$ for which

$$C_1(x+2) < \frac{\varphi(x)}{1-\Phi(x)} < C_2(x+2), \quad \text{if } x \ge -1$$
$$C_1(x+2) < \frac{\varphi(-x)}{\Phi(-x)} < C_2(x+2), \quad \text{if } x \ge -1.$$

b.) There exist some constants $C_1 > 0$ and $C_2 > 0$ such that for all $x > 0$ and $|h| < x+1$

$$e^{-C_1 h(x+2)} < \frac{1-\Phi(x+h)}{1-\Phi(x)} < e^{-C_2 h(x+2)}, \quad \text{if } x \ge 0, \text{ and } h > 0,$$
$$e^{C_2 h(x+2)} < \frac{\Phi(-x+h)}{\Phi(-x)} < e^{C_1 h(x+2)}, \quad \text{if } x \ge 0 \text{ and } h > 0,$$

and (considering the cases $h > 0$ and $h < 0$ separately)

$$e^{C_1 |h|(x+2)} < \frac{1-\Phi(x+h)}{1-\Phi(x)} < e^{C_2 |h|(x+2)}, \quad \text{if } x \ge 0, \text{ and } h < 0,$$
$$e^{-C_2 |h|(x+2)} < \frac{\Phi(-x+h)}{\Phi(-x)} < e^{-C_1 |h|(x+2)}, \quad \text{if } x \ge 0 \text{ and } h < 0.$$

Let η be a random variable with standard normal distribution and $F(x)$ an arbitrary distribution function. It is known (see e.g. problem 7 in the series of problems *The relation between the closeness of probability measures and random variables* that the random variable $\xi = F^{-1}(\Phi(\eta))$ has distribution function $F(x)$, where the inverse function $F^{-1}(x)$ is defined as $F^{-1}(x) = \sup\{u\colon F(u) < x\}$, in the general case. (This is the quantile transform.) In the next problem we give an estimate about the difference of the random variables η and the above constructed ξ if the distribution function $F(x) = F_n(x)$ of the random variable ξ is the normalization of the binomial distribution with parameters n and $p = \frac{1}{2}$, or a little bit more generally we consider such a distribution function where we divide by a number $\dfrac{\sqrt{m}}{2}$ instead of the square root of the variance $\dfrac{\sqrt{n}}{2}$, and the numbers m and n are close to each other.

2.) Let $F_{m,n}(x)$ be the distribution function of a random variable

$$\xi_{m,n} = \frac{2}{\sqrt{m}}\left(\bar{\xi}_n - \frac{n}{2}\right),$$

which is a linear transform of a binomial $B(n, \frac{1}{2})$ distributed random variable $\bar{\xi}_n$. That is, we define $\bar{\xi}_n$ in the following way: Let χ_k, $k = 1, \dots, n$, $P(\chi_k = 0) =$

$P(\chi_k = 1) = \frac{1}{2}$, be independent and identically distributed random variables, and put $\bar{\xi}_n = \sum\limits_{k=1}^{n} \chi_k$. We shall consider the case when $|n-m| \le Bn$ with a sufficiently small number $B > 0$. Let us prove with the help of the above formulated large deviation *Theorem* that the distribution function $F_{n,m}(x)$ satisfies the estimate

$$1 - F_{m,n}(x) = (1-\Phi(x)) \exp\left\{ O\left(\frac{x^3 + \frac{|n-m|}{\sqrt{n}}(x^2+x)+1}{\sqrt{n}} \right) \right\}$$

$$F_{m,n}(-x) = \Phi(-x) \exp\left\{ O\left(\frac{x^3 + \frac{|n-m|}{\sqrt{n}}(x^2+x)+1}{\sqrt{n}} \right) \right\}$$

if $0 \le x \le A\sqrt{n}$ and $|n-m| < Bn$ with sufficiently small constants $A > 0$ and $B > 0$. The $O(\cdot)$ in the above formula is uniform in the variables n, m and x.

Let η be a random variable with standard normal distribution, and define an $F_{m,n}$ distributed random variable $\xi_{m,n} = F_{m,n}^{-1}(\Phi(\eta))$. Prove with the help of the previous estimate that there exist constants $A > 0$ and $K > 0$ and a threshold $n_0 = n_0(B)$ in such a way that for all numbers $n > n_0$

$$|\xi_{m,n} - \eta| < K \frac{1 + \eta^2 + \frac{(m-n)^2}{n}}{\sqrt{n}} \quad \text{on the set } \{|\eta| \le A\sqrt{n}\}.$$

We give a short informal description of the construction which satisfies the *Approximation Theorem*. Let us fix the Brownian bridge $B(t)$. We construct the empirical distribution function $P_n(t)$ and its standardization $Z_n(t) = \sqrt{n}(P_n(t) - t)$, $0 \le t \le 1$, in a recursive way. After the l-th step the random variables $Z_n\left(\frac{k}{2^l}\right)$, $k = 0, 1, \dots, 2^l$, are already defined, and in the $l+1$-th step we construct the random variables $Z_n\left(\frac{2k+1}{2^{l+1}}\right)$, $k = 0, 1, \dots, 2^l - 1$. In the definition of these random variables we have to handle the conditional joint distribution of these random variables under the condition that the previously constructed random variables $Z_n\left(\frac{k}{2^l}\right)$, $k = 0, 1, \dots, 2^l$, take prescribed values. It is more convenient to work with the differences

$$Z_n\left(\frac{2k+1}{2^{l+1}}\right) - Z_n\left(\frac{k}{2^l}\right) \text{ or } Z_n\left(\frac{k+1}{2^l}\right) - Z_n\left(\frac{2k+1}{2^{l+1}}\right)$$

instead of the random variables $Z_n\left(\frac{2k+1}{2^{l+1}}\right)$. We have to describe the conditional joint distribution of these differences under the condition that the values of the random variables $Z_n\left(\frac{k}{2^l}\right)$, $k = 0, 1, \dots, 2^l$, are prescribed. Beside this we investigate the analogous problem about the Brownian distribution. We describe the conditional joint distribution of the random vector whose elements are the differences $B\left(\frac{2k+1}{2^{l+1}}\right)-B\left(\frac{k}{2^l}\right)$ or $B\left(\frac{k+1}{2^l}\right)-B\left(\frac{2k+1}{2^{l+1}}\right)$ under the condition that the values of the random variables $B\left(\frac{k}{2^l}\right)$, $k=1,\dots,2^l$, are prescribed. We shall see that because of the Markov property of the processes $Z_n(t)$ and $B(t)$ the coordinates of the above random vectors are conditionally independent under the conditions we have imposed. We shall construct the above differences of the standardized empirical distribution function by means of a (conditional) quantile transform of the appropriate differences of the Brownian bridge. The word conditional refers to the fact that we shall work with conditional distributions. As a detailed analysis will later show we can guarantee in such a way that the process we shall define has the right multi-dimensional distributions. The content of the expression “conditional quantile transform” will be clearer when the details are worked out.

We shall see that by applying the above sketched argument we have to work with such conditional distributions which can be well approximated by means of the Theorem formulated at the beginning of this text. In particular, the result of problem 2 will be useful. We shall apply the (conditional) quantile transform directly not to the differences of the processes $Z_n(t)$ and $B(t)$ mentioned in the previous paragraph, but we take out from this differences their conditional expectation under the condition that the already defined random variables have their right value. In such a way we shall apply the (conditional) quantile transform between two (conditional) distribution functions whose expected values agree, but whose variances may differ. This is the reason why we considered in problem 2 the approximation of a not necessarily appropriately normalized binomially distributed random variable by a standard normal random variable.

The conditional expectations appearing in this procedure are simple, hence we can get a good bound on the supremum of the absolute value of the random variables

$$Z_n\left(\frac{k+1}{2^l}\right)-B\left(\frac{k+1}{2^l}\right),\quad k=0,1,\dots,2^l,$$

if we apply the above construction. The bound obtained in such a way is not sufficient in itself to prove the *Approximation Theorem.* But it is such an expression which can be bounded well by means of classical methods of the Probability Theory, and by doing this we get the desired result.

The crucial part of the proof of the *Approximation Theorem* will consist of working out the details of the above procedure. To carry it out we introduce some notations.

Let a standardized empirical process $Z_n(t)$ or a Brownian bridge $B(t)$ be given. We shall define some random vectors by means of “successive halving” of the parameter t, $0\le t\le 1$, of these processes together with an appropriate normalization. In the subsequent problems 3 and 4 we formulate the most important properties of these random vectors. These properties will suggest the construction of a process $Z_n(t)$ which satisfies the *Approximation Theorem.*

Put

$$U_{k,l} = 2^{(l+1)/2}\left[B\left(\frac{k}{2^l}\right) - B\left(\frac{k-1}{2^l}\right)\right], \quad k = 1,2,\dots,2^l, \; l = 0,1,2,\dots \tag{1a}$$

$$V_{k,l}(n) = 2^{(l+1)/2}\left[Z_n\left(\frac{k}{2^l}\right) - Z_n\left(\frac{k-1}{2^l}\right)\right], \quad k = 1,2,\dots,2^l, \; l = 0,1,2,\dots, \tag{1b}$$

i.e. we consider an appropriate normalization of the differences of the processes $B(t)$ and $Z_n(t)$ between the diadic points $k2^{-l}$, $k = 0,1,\dots,2^l$. Let us also introduce the σ-algebras

$$\mathcal{F}_l = \mathcal{B}\left\{U_{k,l},\ 1 \le k \le 2^l\right\}, \quad l = 1,2,\dots$$

and

$$\mathcal{G}_l = \mathcal{G}_l(n) = \mathcal{B}\left\{V_{k,l}(n),\ 1 \le k \le 2^l\right\}, \quad l = 1,2,\dots \tag{2}$$

Furthermore, we define the random vectors

$$\begin{aligned} \mathbf{U}_l &= \{U_{k,l},\ k = 1,\dots,2^l\}, \quad l = 1,2,\dots \\ \mathbf{V}_l(n) &= \{V_{k,l}(n),\ k = 1,\dots,2^l\}, \quad l = 1,2,\dots \end{aligned} \tag{3}$$

and

$$\bar{\mathbf{U}}_{l+1} = \{\bar U_{1,l+1},\dots,\bar U_{2^{l+1},l+1}\}, \quad \bar{\mathbf{V}}_{l+1}(n) = \{\bar V_{1,l+1}(n),\dots,\bar V_{2^{l+1},l+1}(n)\},$$

where

$$\bar U_{k,l+1} = U_{k,l+1} - E(U_{k,l+1}|\mathcal{F}_l), \quad \bar V_{k,l+1}(n) = V_{k,l+1}(n) - E(V_{k,l+1}(n)|\mathcal{G}_l(n)) \qquad 1 \le k \le 2^{l+1}, \tag{4}$$

and $l = 0,1,2,\dots$.

In the above definitions we could have chosen a different normalization instead of the normalization $2^{(l+1)/2}$. This normalization is natural, because as the subsequent problems show the conditional expectation of the random variables $\bar U_{k,l}$ and $\bar V_{k,l}(n)$ is zero and their conditional variance is almost one under the conditions we shall consider.

3.) Let us apply the previous notations. Let us first observe that the σ-algebra $\mathcal{G}_l(n)$ consists of the following atoms $B(m_1,\dots,m_{2^l})$:

$$B(m_1,\dots,m_{2^l}) = \left\{\omega\colon V_{k,l}(n)(\omega) = \frac{2^{(l+1)/2}}{\sqrt{n}}[m_k - n2^{-l}],\ k = 1,\dots,2^l\right\},$$

where the numbers m_k, $1 \le k \le 2^l$, are non-negative, and $\sum\limits_{k=1}^{2^l} m_k = n$. (The number m_k agrees with the number of those points in the sample $\zeta_1, \dots, \zeta_n$ determing the process $Z_n(t)$ which fall into the interval $[(k-1)2^{-l}, k2^{-l}]$.) The relation

$$E\left(V_{2k-1,l+1}(n)|\mathcal{G}_l(n)\right) = E\left(V_{2k,l+1}(n)|\mathcal{G}_l(n)\right) = \frac{1}{\sqrt{2}} V_{k,l}(n), \quad k = 1, \dots, 2^l$$

holds. Hence

$$\bar{V}_{2k-1,l+1}(n) = -\bar{V}_{2k,l+1}(n) = V_{2k-1,l+1}(n) - \frac{1}{\sqrt{2}} V_{k,l}(n), \quad k = 1, \dots, 2^l.$$

Let us consider the conditional distribution of the random vector $\bar{\mathbf{V}}_{l+1}(n)$ defined in formula (4) with respect to the σ-algebra $\mathcal{G}_l(n)$. On the atom $B(m_1, \dots, m_{2^l})$ of the σ-algebra $\mathcal{G}_l(n)$ this conditional distribution agrees with the distribution of the random vector

$$X = \{X_k,\ k = 1, \dots, 2^{l+1}\} = \{X_k(m_1, \dots, m_{2^l}),\ k = 1, \dots, 2^{l+1}\}$$

where $X_{2k-1} = -X_{2k}$, $k = 1, 2, \dots, 2^l$, and the random variables X_{2k-1}, $k = 1, \dots, 2^l$, are independent. Furthermore, the distribution of the random variable X_{2k-1} agrees with the distribution of the random variable $\left(\frac{2^{l+2}}{n}\right)^{1/2} (\bar{X}_{2k-1} - E\bar{X}_{2k-1})$, where $\bar{X}_{2k-1}$ has binomial distribution $B(m_k, \frac{1}{2})$ with parameters m_k and $\frac{1}{2}$, i.e.

$$P(\bar{X}_{2k-1} = j) = \binom{m_k}{j} 2^{-m_k}, \quad j = 0, 1, \dots, m_k\ .$$

4.) Let us apply the previous notations. The identity

$$E(U_{2k-1,l+1}|\mathcal{F}_l) = E(U_{2k,l+1}|\mathcal{F}_l) = \frac{1}{\sqrt{2}} U_{k,l} \quad k = 1, \dots, 2^l,$$

holds. Hence $\bar{U}_{2k-1,l+1} = -\bar{U}_{2k,l+1} = U_{2k-1,l+1} - \frac{1}{\sqrt{2}} U_{k,l}$ for all $k = 1, \dots, 2^l$.

The random vector $\bar{\mathbf{U}}_{l+1}$ whose elements are defined in formula (4) is independent of the σ-algebra $\mathcal{F}_l$. Its distribution agrees with the distribution of the random vector $Y_1, Y_2, \dots, Y_{2^{l+1}}$, where $Y_{2k-1} = -Y_{2k}$, $k = 1, \dots, 2^l$, and Y_{2k-1}, $k = 1, \dots, 2^l$, are independent random variables with standard normal distribution.

It is worth mentioning that the conditional distribution of the random variable $\bar{U}_{k,l+1}$ (and in particular its conditional variance) with respect to the σ-algebra $\mathcal{F}_l$ does not depend on this condition. This random variables have such a property, because in this problem jointly Gaussian random variables are considered, and by some standard results of the probability theory the joint conditional distribution of certain coordinates of a Gaussian random vector with respect to the remaining coordinates is Gaussian with such a covariance matrix which does not depend on the value of the random variables appearing in the condition. The conditional distribution and variance of the random variables considered in Problem 3 depend on the values of the random variables appearing in the condition, but this dependence is weak. This remark will not be applied in the subsequent discussion, but it may help to understand the ideas behind the proof.

Now we give the precise construction of the standardized empirical distribution function $Z_n(t)$, $0 \le t \le 1$. Put $Z_n(0) = Z_n(1) = 0$, and let us define the random variables $Z_n\left(\frac{k}{2^l}\right)$, $k = 0, 1, \dots, 2^l$, by recursion with respect to the parameter l. If the l-th step of the construction is already done, i.e. the random variables $Z_n\left(\frac{k}{2^l}\right)$, $k = 0, 1, \dots, 2^l$, are already constructed, then let us define the quantities

$$m_k = m_k(l) = \sqrt{n}\left(Z_n\left(\frac{k}{2^l}\right) - Z_n\left(\frac{k-1}{2^l}\right)\right) + \frac{n}{2^l}, \quad k = 1, \dots, 2^l. \tag{5a}$$

The definition of the quantity $m_k = m_k(l)$ has the following content. It tells us that among the independent, in the interval $[0,1]$ uniformly distributed random variables $\zeta_1, \dots, \zeta_n$ which determine the standardized empirical distribution function $Z_n(t)$ how many random variables take their values in the interval $[(k-1)2^{-l}, k2^{-l}]$. We shall exploit that for prescribed numbers $m_k = m_k(l)$, $1 \le k \le 2^l$, the number of points falling in the interval $[(2k-2)2^{-l-1}, (2k-1)2^{-l-1}]$ is a binomially distributed random variable with parameters m_k and $\frac{1}{2}$. Furthermore, these random variables are conditionally independent for different indices k, under the condition that the values m_k, $1 \le k \le 2^l$, are prescribed. In the construction we define these random variables, or because of some technical reasons their appropriate linear transformation, as the quantile transform of some random variables which are natural linear transforms of the Brownian bridge $B(t)$. With the help of these random variables we can express the random variables $Z_n\left(\frac{2k-1}{2^{l+1}}\right)$, $k = 1, \dots, l$, and thus to carry out the $l+1$-th step of the recursion.

Now we describe the construction by means of explicit formulas. Let $\bar F_m(x)$ denote the binomial $B(m, \frac{1}{2})$ distribution function with parameters m and $\frac{1}{2}$, and put

$$F_{m_k,l}(x) = F_{m_k,l,n}(x) = \bar F_{m_k}\left(\frac{\sqrt{n}x}{2^{(l+2)/2}} + \frac{m_k}{2}\right) \tag{5a$'$}$$

with the number $m_k = m_k(l)$ defined in formula (5a). This means in particular that the distribution function $F_{m_k,l}(x)$ is the "almost standardization" of the distribution function $\bar F_{m_k}(x)$. Indeed, $F_{m_k,l}(x)$ is the distribution function of such a transformation of a $B(m_k, \frac{1}{2})$ distributed random variable in which we take from this random variable its expected value $\frac{m_k}{2}$, but divide by $\frac{\sqrt{n}}{2\cdot 2^{l/2}}$ instead of the square root of the variance $\frac{1}{2}\sqrt{m_k}$. Put

$$\bar V_{2k-1,l+1}(n) = F^{-1}_{m_k,l}\left(\Phi(\bar U_{2k-1,l+1})\right), \quad k = 1, \dots, 2^l, \tag{5b}$$

where $F^{-1}(x)=\sup\{u\colon F(u)<x\}$, the number $m_k=m_k(l)$ and distribution function $F_{m_k,l}(x)$ are defined in formulas (5a) and (5a′), and the random variable $\bar U_{2k-1,l+1})$ is defined in formulas (1)–(4) by means of the Brownian bridge we want to approximate by a standardized empirical function. This means that the random variable $\bar V_{2k-1,l+1}(n)$ is the quantile transform of the random variable $\bar U_{2k-1,l+1}$ which is a functional of the originally fixed Brownian bridge $B(t)$, $0\le t\le 1$, and has standard normal distribution. We have defined the distribution function $F_{m_k,l}(x)$ in the above quantile transform in such a way as the properties of the distributions of the random variables $\bar V_{2k-1,l+1}(n)$ defined in formulas (1)—(4) suggest.

Let us also define the random variables

$$V_{2k-1,l+1}(n)=\bar V_{2k-1,l+1}(n)+\frac{V_{k,l}(n)}{\sqrt2} \qquad (5c)$$
$$\left(=\bar V_{2k-1,l+1}(n)+E\left(V_{2k-1,l+1}(n)|\mathcal G_l(n)\right),\right)\quad k=1,\dots,2^l,$$

and

$$Z_n\left(\frac{2k-1}{2^{(l+1)}}\right)=Z_n\left(\frac{2k-2}{2^{(l+1)}}\right)+2^{-(l+2)/2}V_{2k-1,l+1}(n),\quad k=1,\dots,2^l. \qquad (5d)$$

These definitions follow the following line of argument. First we construct the random variables $\bar V_{2k-1,l+1}(n)$ by means of quantile transform. Then we define the remaining random variables with the help of these random variables $\bar V_{2k-1,l+1}(n)$ by "inverting" formulas (1)—(4). In this "inversion" we also apply the results of problem 3. This principle also suggest the following formulas. Put

$$V_{2k,l+1}(n)=\sqrt2V_{k,l}(n)-V_{2k-1,l+1}(n) \qquad (5e)$$

and

$$\bar V_{2k,l+1}(n)=-\bar V_{2k-1,l+1}(n), \qquad (5f)$$

because formulas (1)—(4) and the results of problem 3 suggest such a definition. In the subsequent problems we complete the construction of the sequence of random variables $\zeta_1,\dots,\zeta_n$ which should satisfy should *Approximation Theorem*, and show that they are independent random variables with uniform distribution in the interval $[0,1]$.

5a.) Let us fix an integer $L>0$, and let us construct for all constants $l=0,1,\dots,L-1$ the random variables $Z_n(k2^{-(l+1)})$, $V_{k,l+1}(n)$, $\bar V_{k,l+1}(n)$, $1\le k\le 2^{(l+1)}$, through formulas (5a)–(5f) by induction with respect to the parameter l. (We define $Z_n(0)\equiv Z_n(1)\equiv 0$, $V_{0,1}(n)\equiv 0$ and $m_0=n$ which relations are needed to start this construction for $l=0$). The distribution of the random vector $Z_n(k2^{-L})$, $1\le k\le 2^L$, defined in such a way agrees with the joint distribution of the values of a normalized empirical distribution function in the coordinate points $t=k2^{-L}$, $k=1,\dots,2^L$. The random vectors $V_{k,l}(n)$ and $\bar V_{k,l}(n)$ constructed in such a way and the process $Z_n(t)$ (more precisely its restriction to the points a $k2^{-L}$ which we have constructed through this procedure) satisfy the properties (1)—(4). More precisely they satisfy formula (1b) for $l \le L$ and that part of formula (4) which contains the random variables V and $\bar V$ with indices $l \le L-1$. Furthermore, the relation $\mathcal{G}_l(n) \subset \mathcal{F}_l$ holds for all numbers $l \le L$.

5b.) Let us consider the random variables $Z_n(k2^{-L})$, $0 \le k \le 2^L$, defined in the previous problem, ($Z_n(0) \equiv 0$), and put $m_k = \sqrt{n}\left(Z_n(k2^{-L}) - Z_n((k-1)2^{-L})\right) + n2^{-L}$, $1 \le k \le 2^L$. For all numbers $k = 1, \dots, 2^L$ let us construct a sequence $\zeta_1^{(k)}, \dots, \zeta_{m_k}^{(k)}$ of independent and in the interval $[(k-1)2^{-L}, k2^{-L}]$ uniformly distributed random variables consisting of m_k terms. For different numbers k let these sequences $\zeta_1^{(k)}, \dots, \zeta_{m_k}^{(k)}$ be independent of each other.

Let us consider the union of these random sequences, and put the elements of this unified sequence in an increasing order. The random sequence $0 \le \zeta_1^* \le \zeta_2^* \le \cdots \le \zeta_n^*$ obtained in such a way is an ordered sample made from n independent and in the interval $[0,1]$ uniformly distributed random variables, i.e. its distribution agrees with the distribution of a sequence which we obtain by putting the elements of a sequence $\zeta_1, \dots, \zeta_n$ of independent random variables with uniform distribution on the interval $[0,1]$ into increasing order. If we choose one of the permutations $(\pi(1), \dots, \pi(n))$ of the set $\{1, \dots, n\}$ randomly and independently of the random sequence $0 \le \zeta_1^* \le \zeta_2^* \le \cdots \le \zeta_n^*$, and in such a way that all possible permutations of the set $\{1, \dots, n\}$ are chosen with the same probability $\frac{1}{n!}$, then the coordinates of the random vectors $(\zeta_1, \dots, \zeta_n) = (\zeta_{\pi(1)}^*, \dots, \zeta_{\pi(n)}^*)$ are independent and in the interval $[0,1]$ uniformly distributed random variables.

Let us construct the standardized empirical distribution function $Z_n(t)$ with the help of the above constructed sequence $(\zeta_1, \dots, \zeta_n)$ of independent and in the interval $[0,1]$ uniformly distributed random variables. The values of this standardized empirical distribution functions in the points $t = k2^{-L}$, $0 \le k \le 2^L$, equal the previously defined random variables $Z_n(k2^{-L})$, $1 \le k \le 2^L$.

We want to show that the originally given Brownian bridge $B(t)$ and standardized empirical distribution function $Z_n(t)$ made from the sequence of independent and in the interval $[0,1]$ uniformly distributed random variables $\zeta_1, \dots, \zeta_n$ constructed in problems 5a and 5b satisfy the *Approximation Theorem* if the number $L = L(n)$ of halving in problems 5a and 5b is sufficiently large. For instance $L = L(n) = n$ is an appropriate choice. Furthermore, we assume that $n \ge n_0$ with a sufficiently large fix number n_0. In order to prove the *Approximation Theorem* we shall estimate the probabilities

$$P\left(\sup_{1 \le k \le 2^l} \sqrt{n}|Z_n(k2^{-l}) - B(k2^{-l})| > \frac{x}{2}\right) \tag{6a}$$

$$P\left(\sqrt{n} \sup_{(k-1)2^{-l} \le t \le k2^{-l}} \left| Z_n(t) - Z_n\left(\frac{k-1}{2^l}\right)\right| > \frac{x}{4}\right) \tag{6b}$$

and

$$P\left(\sqrt{n}\sup_{(k-1)2^{-l}\le t\le k2^{-l}}\left|B(t)-B\left(\frac{k-1}{2^l}\right)\right|>\frac{x}{4}\right) \tag{6c}$$

for all numbers $x > C_0\log n$, with appropriate constants $C_0 > 0$, $l = l(x)$ and all numbers $1 \le k \le 2^l$. We shall deduce *the Approximation Theorem* from a good bound on the above probability together with an appropriate choice of the parameter $l = l(x)$ in them.

The hardest and most important task is to estimate the probability in formula (6a). Let us also remark that from the above expressions only formula (6a) depends on the construction, i.e. on the joint distribution of the processes $Z_n(t)$ and $B(t)$. To give a good estimate on it first we formulate the following problem 6, whose solution follows from the result of problem 2 and the structure of the above described construction of the process $Z_n(t)$.

6.) Let $U_{k,l}$, $V_{k,l}(n)$ $\bar U_{k,l}$ and $\bar V_{k,l}(n)$ be the random variables defined by means of formulas (1)—(4) through a Brownian bridge $B(t)$ and the previously constructed empirical process $Z_n(t)$. Let us consider such a number l for which $l \le L$, where L is the number appearing in the construction of the standardized empirical number $Z_n(t)$, "the number of halving". Let us show that there exist some constants $K > 0$ and $A > 0$ such that the inequalities

$$\begin{aligned}2^{-(l+2)/2}\left|\bar U_{2k-1,l+1}-\bar V_{2k-1,l+1}(n)\right| &= 2^{-(l+2)/2}\left|\bar U_{2k,l+1}-\bar V_{2k,l+1}(n)\right|\\ &< \frac{K}{\sqrt{n}}\left(\bar U^2_{2k-1,l+1}+V^2_{k,l}(n)+1\right)\\ &= \frac{K}{\sqrt{n}}\left(\bar U^2_{2k,l+1}+V^2_{k,l}(n)+1\right),\end{aligned} \tag{7a}$$

and

$$\begin{aligned}&\max\left(2^{-(l+2)/2}\left|U_{2k-1,l+1}-V_{2k-1,l+1}(n)\right|,\ 2^{-(l+2)/2}\left|U_{2k,l+1}-V_{2k,l+1}(n)\right|\right)\\ &\quad< \frac{K}{\sqrt{n}}\left(\bar U^2_{2k-1,l+1}+V^2_{k,l}(n)+1\right)+\frac{2^{-(l+1)/2}}{2}\left|U_{k,l}-V_{k,l}(n)\right|\\ &\quad= \frac{K}{\sqrt{n}}\left(\bar U^2_{2k,l+1}+V^2_{k,l}(n)+1\right)+\frac{2^{-(l+1)/2}}{2}\left|U_{k,l}-V_{k,l}(n)\right|\end{aligned} \tag{7b}$$

hold for all $1 \le k \le 2^l$ if $|\bar U_{2k-1,l+1}| < A\sqrt{n}2^{-l/2}$ and $|V_{k,l}(n)| < A\sqrt{n}2^{-l/2}$.

The result of problem 6 gives a good estimate on the goodness of the quantile transform approximation applied in our construction. Let us remark that the distribution of the term $(\bar U^2_{2k,l+1} + V^2_{k,l}(n) + 1)$ at the right hand side of the expressions investigated in problem 6 can be well bounded, and its distribution is exponentially decreasing at infinity. The estimates given in problem 6 are more useful for us than a sharp estimate on the distribution of the expression at the left-hand side of these expressions. The reason for it is that the differences of the form $Z_n(k2^{-l}) - B(k2^{-l})$ which we want to study can be expressed as appropriate linear combinations of such expressions whose absolute values are bounded in problem 6. We get a better bound on these linear combinations and as a consequence on the distribution of the expression bounded in the *Approximation Theorem* if we exploit not only the smallness of the terms in these linear combinations but also the cancellations among them.

For the sake of further investigations we define the order of a diadic rational number. We say that the diadic order of the number $t = k2^{-l}$, $0 \le t \le 1$, is l, if the number k in the above presentation is odd. Let us consider a number $t = k2^{-l}$ whose diadic order is l. We claim that there exists a sequence of intervals I_j, $0 \le j \le l$, in such a way that $I_0 = [0,1]$, $I_0 \supset I_1 \supset \cdots \supset I_l \ni t$, and $I_j = [u_j, v_j] = [u_j(t), v_j(t)] = [k_j(t)2^{-j}, (k_j(t)+1)2^{-j}]$, where $k_j(t)$ is an integer, i.e. I_j is an interval of length 2^{-j}, and its endpoints are diadic rational numbers whose diadic order is not greater than j.

Indeed, put $I_0 = [0,1]$, and let I_1 be those interval between the intervals $[0, \frac{1}{2}]$ and $[\frac{1}{2}, 1]$ which contains the point t. If the interval $I_j = [u_j, u_j + 2^{-j}]$ is already defined for some $j < l$, then let I_{j+1} be that interval $[u_j, u_j + 2^{-j-1}]$ between the intervals $[u_j + 2^{-j-1}, u_j + 2^{-j}]$ which contains the point t. For $j+1 < l$ this definition is unique. For $j+1 = l$ both interval could be chosen. For the sake of a definite definition in this case we define the interval I_l as $I_l = [t, t+2^{-l}]$.

First we show that the random variables $Z_n(t)$ and $B(t)$ can be expressed as an appropriate linear combination of the random variables $V_{k,j}(n)$ and $U_{k,j}$ corresponding to the above constructed intervals $I_j = I_j(t)$.

7.) Let $t = k2^{-l}$ be a diadic rational number with diadic order l. Let us consider the above defined intervals $I_j = I_j(t) = [u_j, v_j]$, $0 \le j \le l$, and define the quantities $\varepsilon(j) = \varepsilon(j,t)$, $1 \le j \le l$, by the formula $\varepsilon(j) = 0$, if $u_{j-1} = u_j$, and $\varepsilon(j) = 1$, if $u_j > u_{j-1}$, $1 \le j \le l$. Let us introduce the notation $k_j = u_j 2^j$, $0 \le j \le l$. Furthermore, let $U_{k,l}$, $V_{k,l}(n)$ and $\bar U_{k,l}$, $\bar V_{k,l}(n)$ be the random variables defined in formula (1)—(4). Then

$$
\begin{aligned}
Z_n(t) = Z_n\left(k2^{-l}\right) &= \sum_{j=1}^{l} \varepsilon(j) 2^{-(j+1)/2} \left(\sqrt{2} V_{k_{j-1}+1,j-1}(n) - V_{k_j+1,j}(n)\right) \\
B(t) = B\left(k2^{-l}\right) &= \sum_{j=1}^{l} \varepsilon(j) 2^{-(j+1)/2} \left(\sqrt{2} U_{k_{j-1}+1,j-1} - U_{k_j+1,j}\right).
\end{aligned} \tag{8a}
$$

Furthermore,

$$
\begin{aligned}
&2^{-(j+1)/2} \left|V_{k_j+1,j}(n) - U_{k_j+1,j}\right| \\
&\qquad \le \frac{K}{\sqrt{n}} \left(\sum_{s=0}^{j-1} 2^{-s} \left(\bar U^2_{k_{j-s}+1,j-s} + V^2_{k_{j-s-1}+1,j-s-1}(n) + 1 \right) \right),
\end{aligned} \tag{8b}
$$

and

$$\sqrt{n}|Z_n(t)-B(t)| = \sqrt{n}\left|Z_n(k2^{-l})-B(k(2^{-l})\right| \\ \le 4K\left(\sum_{j=1}^{l}\left(\bar U^2_{k_j+1,j}+V^2_{k_{j-1}+1,j-1}(n)+1\right)\right) \tag{8c}$$

on the set

$$\mathbf{B}=\mathbf{B}(t)=\mathbf{B}(k2^{-l}) \\ =\bigcap_{j=1}^{l}\left(\left\{\omega\colon |\bar U_{2k_{j-1}+1,j}(\omega)|<\frac{A\sqrt{n}}{2^{j/2}}\right\}\cap\left\{\omega\colon |V_{k_{j-1}+1,j-1}(n)(\omega)|<\frac{A\sqrt{n}}{2^{j/2}}\right\}\right),$$

where the constants $K>0$ and $A>0$ agree with those introduced in problem 6.

We shall show that the probability in formula (6a) will be bounded in the following way.

$$P\left(\sup_{1\le k\le 2^n}\sqrt{n}\left|Z_n\left(k2^{-l}\right)-B\left(k2^{-l}\right)\right|>\frac{x}{2}\right)<e^{-Dx} \tag{9}$$

with an appropriate number $D>0$, if $C_0\log n\le x\le C^{-1}n$ and $2^{-l}\ge Cxn^{-1}$,

where $C_0>0$ and $C>0$ are sufficiently large positive numbers $n\ge n_0$, and $n_0=n_0(C_0,C)$ is an appropriate threshold index.

The restriction $x>C_0\log n$ in the estimate (9) makes no problem in the proof of the *Approximation Theorem* because in this result only the case $x\ge$ const. $\log n$ has to be considered. Neither the condition $x\le C^{-1}n$ causes a big problem, because the case $x\ge C^{-1}n$ can be simply handled in the proof of the *Approximation Theorem.* The inequality $2^{-l}\le Cxn^{-1}$ imposed for the number $l=l(x)$ tells us how dense subset can be taken in the estimate (9). It will turn out that this subset is sufficiently dense for our purposes, and the *Approximation Theorem* will follow from the estimate (9) and a good bound on the expressions in formulas (6b) and (6c).

Formula (9) will be deduced from formula (8c) and the results of the subsequent problems 8–12. These problems contain the conditions imposed in formula (9).

8.) The joint distribution of the random variables $\bar U_{2k_{j-1}+1,j}$, $1\le j\le l$ appearing in inequality (8c) and the joint distribution of the random variables $\bar U_{1,j}$, $1\le j\le l$, agree. Similarly, the joint distribution of the random variables $V_{k_{j-1}+1,j-1}(n)$, $1\le j\le l$, and $V_{1,j-1}(n)$, $1\le j\le l$, $1\le j\le l$, agree. Furthermore, the random variables $\bar U_{1,j}$, $1\le j\le l$ are independent, and they have standard normal distribution.

Let the inequalities $C_0\log n\le x\le C^{-1}n$ and $2^{-l}\ge C\frac{x}{n}$ hold with some sufficiently large constants $C_0>0$ and $C>0$. Then the probability of the set $\mathbf{B}$ introduced in problem 7 satisfies the inequality $1-P(\mathbf{B})\le e^{-D_1x}$ with an appropriate constant

$D_1 = D_1(C) > 0$. If the above considered constant $C > 0$ is sufficiently large and $n \ge n_0$, where n_0 is an appropriate threshold, then

$$P\left(18K\sum_{j=1}^{l}\bar U_{1,j}^2 > x\right) \le e^{-D_2 x}$$

with an appropriate number $D_2 > 0$. (In this formula the same constant K appears as in the problems 6 and 7.)

We remark that the last estimate (if we are not interested in the choice of the constants in them) is sharp. Indeed, even a single term of this sum, being the square of a random variable with standard normal distribution, takes a value larger than $x > 0$ with probability larger than $e^{-\text{const.}\, x}$. This means that disregarding the constant in the exponent we can get as good an estimate for the probability that this sum consisting of l terms is greater than x, as for the probability of the event that a single term of the sum is larger than this bound. Such an estimate holds if the number of terms are chosen so that the expected value of the sum is smaller than αx with some number $0 < \alpha < 1$.

To prove formula (9) we shall show that

$$\begin{aligned} P\left(18K\sum_{j=1}^{l} V_{k_{j-1}+1,j-1}(n)^2 > x\right) &= P\left(18K\sum_{j=1}^{l} V_{1,j-1}(n)^2 > x\right) \\ &= P\left(18K\sum_{j=1}^{l} 2^j Z_n\left(\frac{1}{2^{j-1}}\right)^2 > x\right) \le e^{-D_3 x} \end{aligned} \tag{10}$$

with some appropriate constant $D_3 > 0$, if $C_0 \log n \le x \le C^{-1} n$ and $2^{-l} \ge Cxn^{-1}$ with appropriate constants $C_0 > 0$ and $C > 0$. (The first two identities in formula (10) follow from the result of problem 8 and the definition of the random variables appearing in formula (10).) First we consider the following problem.

9.) Let us deduce formula (9) from the results of problems 7 and 8 and the estimate in formula (10).

Let us remark that formula (10) is similar to the last estimate of problem 8. Furthermore, the standardized empirical process $Z_n(t)$ behaves similarly to the Brownian bridge $B(t)$. It is not difficult to prove such an analog of formula (10) where the random variables $V_{k_{j-1}+1,j-1}(n)$ are replaced by the random variables $U_{k_{j-1}+1,j-1}$. This suggests that estimate (10) should also hold. The main technical difficulty in its proof arises because the standardized empirical distribution function $Z_n(t)$ is not a process with independent increment. Hence the most powerful methods of probability theory which are mainly appropriate for the investigation of processes with independent increments do not supply a simply way to prove formula (10). We shall overcome this difficulty by applying a classical method the so-called Poissonian approximation.

First we describe the Poissonian approximation we shall apply and explain how it helps to simplify the proof of formula (10). Let ζ_k, $k=1,2,\dots$, be a sequence of independent and in the interval $[0,1]$ uniformly distributed random variables, and let κ_n be a Poisson distributed random variable with parameter n, i.e. let $P(\kappa_n=j)=\frac{n^j}{j!}e^{-n}$, $j=0,1,2,\dots$, which is independent of the sequence ζ_k, $k=1,2,\dots$. Define the stochastic processes

$$Z_n(t)=\frac{1}{\sqrt{n}}\left(\sum_{k=1}^{n} I(\{\zeta_k\le t\})-nt\right) \tag{11a}$$

$$X_n(t)=\frac{1}{\sqrt{n}}\left(\sum_{k=1}^{\kappa_n} I(\{\zeta_k\le t\})-nt\right) \tag{11b}$$

and

$$Y_n(t)=\frac{1}{\sqrt{n}}\sum_{k=n}^{\kappa_n} I(\{\zeta_k\le t\}),\qquad 0\le t\le 1, \tag{11c}$$

where $I(A)$ denotes the indicator set of the set A, and in the case $\kappa_n<n$ the sum which defines the random process $Y_n(t)$ is meant as $\sum\limits_{k=n}^{\kappa_n}=-\sum\limits_{k=\kappa_n}^{n}$. Then $Z_n(t)$, $0\le t\le 1$, is a standardized empirical distribution function, and $Z_n(t)=X_n(t)-Y_n(t)$. Hence it is not difficult to show that

$$\begin{aligned}P\left(18K\sum_{j=1}^{l}2^jZ_n\left(\frac{1}{2^{j-1}}\right)^2>x\right)\le P\left(72K\sum_{j=1}^{l}2^jX_n\left(\frac{1}{2^{j-1}}\right)^2>x\right)\\+P\left(72K\sum_{j=1}^{l}2^jY_n\left(\frac{1}{2^{j-1}}\right)^2>x\right).\end{aligned} \tag{12}$$

The proof of the inequality (12) will be part of the problems 10 and 11. We shall be able to bound the second term at the right-hand side of inequality (12), because the sum defining the random variable $Y_n(\cdot)$ contains relatively few, $|\kappa_n-n|$ terms. Hence it can be bounded with the help of relatively weak estimations. This will be done in problems 10 and 11.

The estimation of the first term at the right-hand side of inequality is relatively simple, because $X_n(t)$ is a standardized Poisson process with parameter n, i.e. for arbitrary constants $0\le t_0<t_1<\cdots<t_k\le 1$ the random variables $X_n(t_j)-X_n(t_{j-1})$, $1\le j\le k$, are independent, and the random variable $\sqrt{n}\,(X_n(t_j)-X_n(t_{j-1}))+n(t_j-t_{j-1})$ is Poisson distributed with parameter $n(t_j-t_{j-1})$. It is a well known result of probability theory that if the above defined process $X_n(t)$ is considered, then the process $\sqrt{n}X_n(t)+nt$, $0\le t\le 1$, is a Poisson process with parameter n. (The proof of this

result is also contained in the solution of problem 2 in the series of problems *Poisson processes* on my homepage, but it exists now only in Hungarian.)

10.) Let us prove formula (12). Let us also show that if κ_n is a Poisson distributed random variables with parameter n, then for all numbers y such that $0 \le y \le B_1\sqrt{n}$ with some constant $B_1 > 0$ there exists a constant $B_2 = B_2(B_1) > 0$ such that $P(|\kappa_n - n| \ge y) \le e^{-B_2 y^2/n}$.

Let us fix some positive real number $x > 0$ and positive integer l. Let ζ_k, $k = 1, 2, \dots$, be a sequence of independent and in the interval $[0, 1]$ uniformly distributed random variables, and let us define the following stochastic processes $\bar{Y}_{n,m}(t)$ and random variables ξ_k:

$$\bar{Y}_{n,m}(t) = \frac{1}{\sqrt{n}} \sum_{k=1}^{m} I(\{\zeta_k \le t\}), \quad m = 0, 1, 2, \dots,$$

$$\xi_k = \xi_{k,l} = \sum_{j=1}^{l} 2^{j/2} I\left(\zeta_k < \frac{1}{2^{j-1}}\right), \quad k = 1, 2, \dots.$$

Then for arbitrary an constant $B > 0$ the following relation holds:

$$\begin{aligned} P\left(72K\sum_{j=1}^{l} 2^j Y_n\left(\frac{1}{2^{j-1}}\right)^2 > x\right) &\le P\left(|\kappa_n - n| > B\sqrt{nx}\right) \\ &\quad + P\left(\sqrt{72K}\sum_{j=1}^{l} 2^{j/2}\bar{Y}_{n,B\sqrt{nx}}\left(\frac{1}{2^{j-1}}\right) > \sqrt{x}\right) \qquad (13) \\ &= P\left(|\kappa_n - n| > B\sqrt{nx}\right) + P\left(\frac{\sqrt{72K}}{\sqrt{n}} \sum_{k=1}^{B\sqrt{nx}} \xi_k > \sqrt{x}\right). \end{aligned}$$

The random variables $\xi_k = \xi_{k,l}$, $k = 1, 2, \dots$, defined in this problem are independent and identically distributed.

We want to give an in the infinity exponentially decreasing bound for the second term at the right-hand side of formula (12). To do this it is enough to give a good bound on the expression at the right-hand side of formula (13). To do this we have to bound the probability of such an event that the sum of certain independent random variables is greater than a given number. The probability theory has standard methods for the investigation of such problems, and they can be applied also in the present case.

If we want to give a good bound on the probability that a random variable S takes a value greater than $A = A(x)$ with some number A depending on x, more explicitly if we want to give an upper bound of the form $e^{-\text{const.}\, x}$ for the probability of this event, then the following argument is often useful. The inequality $P(S > A) = P\left(\frac{x}{A}S > x\right) \le e^{-x} E\exp\{\frac{x}{A}S\}$ holds. If we can show that $E\exp\{\frac{x}{A}S\} \le e^{\alpha x}$ with some constant $\alpha < 1$, then we get the desired estimate. In the next problem we want to show that this method also works in the case we are interested in.

11.) Let us apply the notations introduced in the previous problem. Let us consider such real constant $x > 0$ and positive integer $l > 0$ which satisfy the inequalities $C_0 \log n < x < C^{-1} n$ and $2^{-l} > Cxn^{-1}$ with some fixed constants $C > 0$ and $C_1 > 0$. Let us show that in this case there exists a constant $\bar{K} = \bar{K}(C) > 0$ such that for all sufficiently large n

$$E \exp\left\{\frac{\sqrt{x}}{\sqrt{n}} \xi_k\right\} \le 1 + \frac{\bar{K}\sqrt{x}}{\sqrt{n}}.$$

Let us prove with the help of the above estimate and the result of problem 10 (by choosing the constant $B > 0$ in it sufficiently small), that for such constants x and l which satisfy the above conditions the inequality

$$P\left(72K \sum_{j=1}^{l} 2^j Y_n \left(\frac{1}{2^{j-1}}\right)^2 > x\right) < e^{-D_4 x} \tag{14}$$

holds with a sufficiently small constant $D_4 > 0$ for all sufficiently large n.

We also want to give a good bound on the distribution of the first term at the right-hand side of formula (12). The standardized Poisson process appearing here is a process with independent increments. But the terms in the sum we have to investigate, the random variables $X_n\left(2^{-j-1}\right)^2$ are not independent, since these terms are the increments of the process $X_n(t)$ in the overlapping intervals $[0, 2^{-j-1}]$. But it is relatively simple to overcome this difficulty, if we write the random variable $X_n\left(2^{-j-1}\right)$ as the sum of the increment of the process $X_n(t)$ on appropriate disjoint intervals, then we bound the square of the sums got in this way by the Cauchy–Schwarz inequality and sum up these inequalities. In such a way we can write that

$$\begin{aligned} 2^j X_n\left(\frac{1}{2^{j-1}}\right)^2 &= 2^j \left(\sum_{k=j}^{l} 2^{-(k-j)/4} \cdot 2^{(k-j)/4} \left[X_n\left(\frac{1}{2^{k-1}}\right) - X_n\left(\frac{1}{2^k}\right)\right]\right. \\ &\qquad\qquad \left. + 2^{-(l-j)/4} \cdot 2^{(l-j)/4} X_n\left(\frac{1}{2^l}\right)\right)^2 \\ &\le 2^j B \left(\sum_{k=j}^{l} 2^{(k-j)/2} \left[X_n\left(\frac{1}{2^{k-1}}\right) - X_n\left(\frac{1}{2^k}\right)\right]^2 + 2^{(l-j)/2} X_n^2\left(\frac{1}{2^l}\right)\right) \end{aligned} \tag{15}$$

for all numbers $j = 1, \dots, l$ with the constant $B = B_j = \sum\limits_{k=j}^{l} 2^{-(k-j)/2} + 2^{-(l-j)/2} < 5$.

We want to show with the help of the inequality (15) that the estimation of the first term in formula (12) can be reduced to the estimation of sums of appropriate linear combination of squares of independent standardized Poisson distributed random variables. The problem which arises in such a way is very similar to the last estimate in problem 8, and it can be solved similarly. However, a new difficulty appears when handling this problem, because the square of a standardized Poisson distributed random variable — unlike the square of a normally distributed random variable — has no exponential moments. This difficulty can be overcome by a natural truncation of the terms in the sum under consideration. The contribution of the terms with too large values have only negligible contribution to the sum, and the truncated random variables have finite exponential moments which can be well bounded. Let us also remark that here we did not really exploited that we work with Poissonian distributed random variables. In the case when the approximation of partial sums of independent random variables with Brownian motion is investigated, then a similar problem discussed appears, but in that problem sums of independent random variables play the role of the Poisson process. It can be solved in the same way because sums of independent random variables with exponential moments have a similarly good Gaussian approximation as a Poissonian random variable with a large parameter.

12.) Let us show that

$$
\begin{aligned}
&P\left(72K\sum_{j=1}^{l}2^jX_n\left(\frac{1}{2^{j-1}}\right)^2>x\right)\\
&\qquad\le P\left(1500K\left(\sum_{k=1}^{l}\frac{2^k}{n}(\eta_k-E\eta_k)^2+\frac{2^l}{n}(\eta_{l+1}-E\eta_{l+1})^2\right)>x\right),
\end{aligned}
\tag{16}
$$

where η_k, $k=1,\dots,l+1$, are independent random variables, η_k has Poissonian distribution with parameter $\lambda_k=\lambda_{k,n}=n2^{-k}$, if $1\le k\le l$, and η_{l+1} has Poissonian distribution with parameter $\lambda_{l+1}=\lambda_{l+1,n}=\lambda_l$.

Let us define the truncation $\bar\eta_k=\bar\eta_k(n)$ of the random variables η_k, $1\le k\le l+1$ in the following way:

$$
\bar\eta_k=\begin{cases}\eta_k-E\eta_k & \text{if } |\eta_k-E\eta_j|<n2^{-k}\\ 0 & \text{if } |\eta_k-E\eta_k|\ge n2^{-k}\end{cases}\qquad k=1,\dots,l+1.
$$

Then the following inequalities hold: $P(|\eta_k-E\eta_k|>u)\le 2\exp\left\{-\dfrac{u^2}{8n2^{-k}}\right\}$, if $u<n2^{-k}$, in particular $P(|\eta_k-E\eta_k|\ge n2^{-k})\le 2\exp\left\{-\dfrac{n}{2^{(k+3)}}\right\}$. Furthermore, $E\exp\left\{\dfrac{2^{k-4}}{n}\bar\eta_k^2\right\}\le B$ with appropriate constant $B>0$ (independent of the parameter n) for all $1\le k\le l+1$.

Let us consider a real number $x>0$ and an integer $l>0$ which satisfy the inequalities $C_0\log n<x<C^{-1}n$ and $2^{-l}>Cxn^{-1}$ with appropriate constants $C>0$ and

$C_0 > 0$, and let the relation $n \ge n_0$ hold, where $n_0 = n_0(C, C_0)$ is an appropriate threshold. Let us show in this case that

$$P\left(72K\sum_{j=1}^{l} 2^j X_n\left(\frac{1}{2^{j-1}}\right)^2 > x\right) \le e^{-D_5 x} \tag{17}$$

with an appropriate constant $D_5 > 0$. Let us show that formulas (10) and (9) follow from the already proved inequalities.

To carry out the proof of the *Approximation Theorem* we still need a good estimate on the probabilities in formulas (6b) and (6c). First we formulate a *Statement* which we shall prove and which supplies great help in bounding the probabilities in formulas (6b) and (6c).

Statement: *Let $B(t)$, $0 \le t \le 1$, be a Brownian bridge, and $Z_n(t)$, $0 \le t \le 1$, a standardized empirical distribution function made from n independent and in the interval $[0,1]$ uniformly distributed random variables. Let us fix a real number $0 \le y \le n$. Then for all real numbers $L > 0$ there exist such positive constant $\alpha = \alpha(L) > 0$ and threshold index $n_0 = n_0(L)$ for which the processes $B(t)$ and $Z_n(t)$ satisfy the following inequalities:*

$$P\left(\sup_{0\le t< L\frac{y}{n}} \sqrt{n}|B(t)| > y\right) \le 2e^{-\alpha y}, \tag{18a}$$

$$P\left(\sup_{0\le t< L\frac{y}{n}} \sqrt{n}|Z_n(t)| > y\right) \le 2e^{-\alpha y}, \tag{18b}$$

if $n \ge n_0$. The constant $\alpha > 0$ and the threshold index n_0 can be given as the function of the number L i.e. the threshold index n_0 can be given independently of the number y, and the exponent $\alpha > 0$ depends neither on the number y nor the threshold index n_0.

We did not try to determine the optimal constants in formulas (18a) and (18b), since we do not need it. The determination of the optimal constant is sufficiently simpler in formula (18a) since the Brownian bridge is a Gaussian process, and the investigation of such processes is considerably simpler. The heuristic content of (18b) is that the process $Z_n(t)$ for large indices n behaves similarly to a Brownian bridge, hence it satisfies similar estimates. The next lemma which supplies a bound on the distribution of the maximum of a process with independent increment plays an important role in the proof of the *Statement*. We shall give the proof of this lemma in an *Appendix*.

Lemma. *Let $\xi_1, \dots, \xi_n$ be independent random variables, for which $E\xi_k \ge 0$, $Ee^{s\xi_k} = e^{B_k(s)}$ with some fixed $s > 0$ and numbers $B_k(s)$, $k = 1, \dots, n$. Put $S_k = \sum\limits_{j=0}^{k} \xi_j$,*

$k = 1, \dots, n$. $S_k = \sum_{j=0}^{k} \xi_j$, $k = 1, \dots, n$. Then

$$P\left(\sup_{1\le k\le n} S_k > x\right) \le \exp\left\{-sx + \sum_{k=1}^{n} B_k(s)\right\}$$

for all $x > 0$.

Let $X(t)$, $a \le t \le b$, be a stochastic process with independent increments in an interval $[a, b]$ i.e. let us assume that the random variables $X(t_1) - X(a)$, $X(t_2) - X(t_1)$, ..., $X(b) - X(t_k)$ are independent for all numbers k and $a \le t_1 \le t_2 \le \dots \le t_k \le b$. Let us also assume that the trajectories of the process $X(t)$ are continuous or more generally so called cadlag (continue à droite, limite à gauche), i.e. continuous from the right functions which have a left-side limit in all points $a \le t < b$. Let us further assume that the function $m(t) = EX(t)$, $a \le t \le b$, is monotone increasing. Then the inequality

$$P\left(\sup_{a\le t\le b} X(t) - X(a) > x\right) \le e^{-sx} E e^{s(X(b)-X(a))}$$

holds for all numbers $s > 0$. (This formula is meant in such a way that the right-hand side of the inequality is infinity in the case when the expectation $Ee^{s(X(b)-X(a))}$ does not exist.)

Remark: The condition $E\xi_k \ge 0$ or the monotonicity of the function $m(t)$ is required, because this guarantees that the partial sums S_k, $k = 1, 2, \dots, n$, or respectively the process $X(t)$, $a \le t \le b$, has a positive trend.

The probability theory has several results which state that the maximum of the partial sum random variables with non-negative expectation is not much greater than the sum of all random variables. The previous *Lemma* is also such a result. It states that a natural bound on the sum of all random variables also supplies a bound for the maximum of all partial sums.

The previous *Lemma* cannot be applied directly for the proof of the *Statement*, since neither the Brownian bridge nor a standardized empirical process are processes with independent increment. But we can prove the *Statement* with the help of the following observation. If $W(t)$, $0 \le t \le 1$, is a Wiener process, that is it is such a Gaussian process for which $EW(t) = 0$, and $EW(s)W(t) = \min(s, t)$ for all $0 \le s, t \le 1$, then $W(t)$ is a process with independent process (and we may also assume that its trajectories are continuous function), and the stochastic process $B(t) = W(t) - tW(t)$ is a Brownian bridge. The other result of the *Statement*, the estimate (18b) can be proved by means of a Poissonian approximation defined with the help of formulas (11a)—(11c).

13.) Let us prove that modification of inequality (18a) in which the Brownian bridge $B(t)$ is replaced by a Wiener process $W(t)$. Let us also prove that modification of the inequality (18b) in which the process $Z_n(t)$ is replaced by a standardized

Poisson process $X_n(t)$ with parameter n. (We have defined a standardized Poisson process with parameter n in formula (11b).)

14.) Let us prove inequality (18) by means of the result of the previous result and the representation of a Brownian bridge $B(t)$ in the form $B(t)=W(t)-tW(1)$, where $W(t)$ is a Wiener process.

Let us consider the process $Y_n(t)$, $0\le t\le 1$, defined in formula (11c). Let us show that this process satisfies the inequality

$$P\left(\sqrt{n}\sup_{0\le t\le L\frac{y}{n}}|Y_n(t)|\ge y\right)\le 2e^{-\alpha y}$$

for all $L\ge 0$ $n>n_0$ with an appropriate threshold index n_0 and appropriate number $\alpha=\alpha(L)>0$. Let us prove with the help of this inequality and the result of the previous problem the inequality (18b).

It is not difficult to prove the *Approximation Theorem* with the help of the above result. Given some number x such that $C_0\log n<x<2C^{-1}n$ with some appropriate constant $C>1$ let us choose a constant $l=l(x)$ in such a way that $C\frac{x}{n}\le 2^{-l}<2C\frac{x}{n}$. Then the previous results enable us both to give a good bound on the probability $P\left(\sup\limits_{1\le k\le 2^l}\sqrt{n}|Z_n(k2^{-l})-B(k2^{-l})|\ge\frac{x}{2}\right)$ and to estimate the fluctuation of the processes $\sqrt{n}B(t)$ and $\sqrt{n}Z_n(t)$ in the intervals $(k-1)2^{-l}\le t<k2^{-l}$, $1\le k\le 2^l$, assuming that the constant $C>0$ (independently of n) is chosen sufficiently large. In the case $x>C^{-1}n$ the proof of the statement of the *Approximation Theorem* is considerably simpler. In this case the rough estimate $\sqrt{n}|Z_n(t)-B(t)|\le\sqrt{n}(|Z_n(t)|+|B(t)|)$ is also sufficient for our goals.

15.) Prove the *Approximation Theorem* with the help of the previous results.

Some comments about the main result discussed in this paper

The main result of this work is contained in the paper of János Komlós, Péter Major and Gábor Tusnády *An approximation of partial sums of independent RV's and the sample DF. I.* in the journal *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **32** (1975), pp. 111–131. This article does not work out all details of the proof. It also contains an analogous result about the approximation of partial sums of independent random variables by a Wiener process. The two results are based on the same ideas and the above mentioned paper works out only the approximation of the partial sums or independent random variables in detail.

The main difference between the discussion of the above mentioned paper and of this work is that here we also worked out those details which are very natural to expect, but whose precise proof is a little bit inconvenient. Such difficulties arise because we have to work with "almost" but not completely independent random variables. Beside this, I thought that a detailed discussion of certain classical methods like for instance the Poissonian approximation may be interesting and instructive in itself. So I tried to explain the ideas behind some technical steps even if it made the discussion considerably longer.

In a subsequent series of problems I shall also discuss the approximation of partial sums of independent random variables. In that work I shall not work out all details. Instead of it I shall try to explain the basic problems in that subject. Furthermore I try to explain with the help of some examples which details must be investigated especially carefully.

The proof of the *Approximation Theorem* is also contained in some works which appeared after the paper of Komlós, Major and Tusnády. Those who are interested more in this subject can study the works of Gábor Tusnády, Jean Bretagnolle, and Pierre Massart. These papers also investigate the problem how small constants can be chosen in the *Approximation Theorem.* Here we have not dealt with this problem.

It us worth mentioning that in some papers the construction satisfying the *Approximation Theorem* and the prof of the result is based on an expansion with respect to Haar functions. Although in my discussion the Haar functions did not appear, the two discussions do not differ essentially. One can say that in these two approaches the same construction is explained in a different language. I briefly write down the argument of the construction made on the basis of expansion with respect to Haar functions.

Let $\varphi_1(t)$, $\varphi_2(t)$, ... , be an arbitrary system of complete orthogonal functions on the space $L_2([0,1],\lambda)$, where λ denotes the Lebesgue measure, and let $\eta_1,\eta_2,\dots,$ be a sequence of independent random variables with standard normal distribution. By a result of probability theory the stochastic process

$$W(t)=\sum_{l=1}^{\infty}\eta_l\int_0^t\varphi_l(s)\,ds,\quad 0\le t\le 1$$

is a Wiener process. Indeed, the above defined process is a Gaussian process (the infinite sum in this expression is convergent with probability one for all numbers $0\le t\le 1$,

$EW(t)=0$, and the covariance function of the process is

$$EW(s)W(t)=\sum_{l=1}^\infty \int_0^s \varphi_l(u)\,du \int_0^t \varphi_l(u)\,du=\min(s,t) \quad \text{for all numbers } 0\le s,t\le 1.$$

The last identity is a consequence of the Parseval formula by which

$$\min(s,t)=\int_0^1 I_{[0,s]}I_{[0,t]}(u)\,du=\sum_{l=1}^\infty a_lb_l,$$

where $I_{[a,b]}(\cdot)$ is the indicator function of the interval $[a,b]$, and $a_l=\int_0^1 I_{[0,s]}(u)\varphi_l(u)\,du$, $b_l=\int_0^1 I_{[0,t]}(u)\varphi_l(u)\,du$. In a complete proof we also should show that the trajectories of the above defined process are continuous. But we shall apply this representation only in a special case, and we shall not need the statement about the continuity of the trajectories. (On the other hand, in the special case we shall consider the proof of the continuity of the trajectories is relatively simple.)

Let us recall the definition of Haar functions. Let us define the functions $\chi_{k,l}(u)$, $0\le l<\infty$, $1\le k\le 2^l$, on the interval $[0,1]$ by the formulas $\chi_{0,1}(u)\equiv 1$, $\chi_{k,l}(u)=2^{l/2}$, if $(k-1)2^{-l}\le u<(2k-1)2^{-l}$, $\chi_{k,l}(u)=-2^{l/2}$, if $(2k-1)2^{-l}\le u<k2^{-l}$ and $\chi_{k,l}(u)=0$ otherwise, $0\le l<\infty$, $1\le k\le 2^k$. These functions $\chi_{k,l}$ are called the Haar functions. It is not difficult to show that the Haar functions constitute a complete orthogonal system in the space $L_2([0,1],\lambda)$. Hence by the previous result on the construction of Wiener processes

$$W(t)=\sum_{l=0}^\infty\sum_{k=1}^{2^l}\eta_{k,l}\int_0^t \chi_{k,l}(s)\,ds, \quad 0\le t\le 1,$$

where the random variables $\eta_{k,l}$ are independent with standard normal distribution. Furthermore, $B(t)=Wt)-tW(1))=W(t)-\int_0^t \chi_{[0,1]}(u)\,du\,W(1)$ is a Brownian bridge, and $W(1)=\eta_{1,0}$, since in the above representation of the random variable $W(1)$ the coefficient of all other random variables $\eta_{k,l}$ is zero, and the coefficient of the variable $\eta_{1,0}$ is one. Hence the stochastic process

$$B(t)=\sum_{l=1}^\infty\sum_{k=1}^{2^l}\eta_{k,l}\int_0^t \chi_{k,l}(s)\,ds, \quad 0\le t\le 1,$$

is a Brownian bridge. Let us also observe that because the special form of the Haar functions

$$\begin{aligned}\eta_{k,l}&=\int \chi_{k,l}(s)B(s)\,ds\\ &=2^{l/2}\left([B(k2^{-l})-B((2k-1)2^{-l-1})]-[B(2k-1)2^{-l-1})-B((k-1)2^{-l})]\right),\end{aligned}$$

and this means that the random variable $\eta_{k,l}$ agrees with the random variable $\bar U_{k,l}$ introduced in our construction.

On the other hand it can be proved by means of induction with respect the parameter l that if we define the random variables $\bar V_{k,l}(n)$ by means of a standardized empirical distribution $Z_n(t)$ through formulas (1b) and (4) (also applying the results of problem 3 which enables a simple calculation of the conditional expectation in formula (4)) that a new process $\bar Z_n(t)$ defined by formula

$$\bar Z_n(t) = \sum_{l=1}^{\infty} \sum_{k=1}^{2^l} \bar V_{k,l}(n) \int_0^t \chi_{k,l}(s)\,ds, \quad 0 \le t \le 1,$$

and the process $Z_n(t)$ satisfy the relation $\bar Z_n(k2^{-l}) = Z_n(k2^{-l})$ for all numbers $l = 1, 2, \dots,$ $1 \le k \le 2^l$. Hence $\bar Z_n(t) \equiv Z_n(t)$ for all parameters $0 \le t \le 1$. If we can give the (joint) construction of a Brownian bridge $B(t)$ and a standardized empirical distribution function in such a way that in their above representation the Fourier coefficients $\bar U_{k,l}$ and $\bar V_{k,l}(n)$ are close to each other, then we get a construction which satisfies the *Approximation Theorem*. Actually this is the construction we made in this paper explained in a different language. The details can be worked out similarly to the argument of the present series of problems.

2. Solution of the problems

1.) The following inequality is a well-known result of probability theory, and it can be simply proved by means of integration by parts. (Otherwise it is also contained in problem 7 of the series of problems *Random variables with normal distribution.* At present it exists only in Hungarian.)

$$\left(\frac{1}{x}-\frac{1}{x^3}\right)\varphi(x) < 1-\Phi(x) < \frac{1}{x}\varphi(x), \quad \text{for all } x>0.$$

It follows from these inequalities that $C_1(x+2) < \dfrac{\varphi(x)}{1-\Phi(x)} < C_2(x+2)$ if $x \ge 2$ with some constants $C_1 > 0$ and $C_2 > 0$. Furthermore, since both functions $\varphi(x)$ and $1-\Phi(x)$ are separated from zero and infinity if x is in a finite domain, the above inequality holds for all $x > -1$. By the identities $\varphi(-x) = \varphi(x)$ and $\Phi(-x) = 1-\Phi(x)$ the second inequality in part a) of problem 1 is equivalent to the first one.

b.)

$$\log\frac{1-\Phi(x+h)}{1-\Phi(x)} = h\,\frac{d}{dx}\log\left(1-\Phi(x)\right)\bigg|_{x=u} = -h\frac{\varphi(u)}{1-\Phi(u)},$$

where u is an appropriate number in the interval $[x, x+h]$. Hence in the case $|h| < |x|+1$

$$C_1(x+2) < \frac{\varphi(u)}{1-\Phi(u)} < C_2(x+2), \quad \text{if } x > -1$$

by the already proven part of the problem. By substituting this relation to the previous identity we get the first identity of the problem in the case $h > 0$ (and $x > -1$). The second identity follows from this one by the relation $\Phi(-u) = 1-\Phi(u)$. The case $h < 0$ can be similarly handled or it can be reduced to the case $h > 0$.

2. First we prove the asymptotic relation for the distribution function $F_{m,n}(x)$.

By the *(large deviation) Theorem* formulated at the beginning of this series of problems and the result of problem 1

$$\begin{aligned}
1-F_{m,n}(x) &= \left(1-\Phi\left(\sqrt{\frac{n}{m}}x\right)\right)\exp\left\{O\left(\frac{x^3+1}{\sqrt{n}}\right)\right\} \\
&= (1-\Phi(x))\exp\left\{O\left(\left|\sqrt{\frac{n}{m}}-1\right|(x^2+2x)+\frac{x^3+1}{\sqrt{n}}\right)\right\} \\
&= (1-\Phi(x))\exp\left\{O\left(\frac{x^3+\frac{|n-m|}{\sqrt{n}}(x^2+x)+1}{\sqrt{n}}\right)\right\}
\end{aligned}$$

if $0 \le x \le \bar{A}\sqrt{n}x$, because $\left|\sqrt{\frac{n}{m}} - 1\right| (x^2 + 2x) \le \text{const.} \frac{|n-m|}{n}(x^2 + 2x)$ if the conditions of problem 2 hold. The other inequality can be proved similarly.

We prove with the help of the already proved inequality that there exists such a constant $K > 0$ for which

$$\begin{aligned} 1 - F_{m,n}(x + h(x)) \le 1 - \Phi(x) \le 1 - F_{m,n}(x - h(x)), \\ \text{with } h(x) = h_{m,n}(x) = K\frac{x^2 + \frac{|n-m|}{\sqrt{n}}(|x|+1) + 1}{\sqrt{n}}, \end{aligned} \tag{2.1}$$

if $0 \le x \le \bar{A}\sqrt{n}$ with some appropriate $\bar{A} > 0$. Indeed, in this case

$$\begin{aligned} \log \frac{1 - F_{m,n}(x + h(x))}{1 - \Phi(x)} &= \log \frac{1 - F_{m,n}(x + h(x))}{1 - \Phi(x + h(x))} + \log \frac{1 - \Phi(x + h(x))}{1 - \Phi(x)} \\ &\le \frac{K_1}{\sqrt{n}} \left((x + h(x))^3 + \frac{|n-m|}{\sqrt{n}} \left((x + h(x))^2 + x + h(x) \right) + 1 \right) - C_1 h(x)(x+2) \end{aligned}$$

with appropriate constants $K_1 > 0$ and $C_1 > 0$ if the inequalities $|x + h(x)| < A\sqrt{n}$ and $|h(x)| \le x + 1$ hold, since if these conditions are satisfied then the results of problem 1 and the already proven part of problem 2 are applicable. In this formula the constant K_1 is the constant which can be written in the $O(\cdot)$ expression of the already proved part of problem 2, and the C_1 is the same constant which appears in part b) of problem 1. We claim that one can choose some constants $\bar{A} > 0$, $B > 0$ and $K > 0$ in such a way that under the conditions $x < \bar{A}\sqrt{n}$, $|n-m| < Bn$, $n \ge n_0$, where n_0 is an appropriate threshold index, the following inequalities hold: $|x + h_{m,n}(x)| < A\sqrt{n}$, $|h(x)| \le x + 1$, and

$$\frac{K_1}{\sqrt{n}} \left((x + h(x))^3 + \frac{|n-m|}{\sqrt{n}} \left((x + h(x))^2 + x + h(x) \right) + 1 \right) - C_1 h(x)(x+2) \le 0.$$

These inequalities and the previous estimates imply that $1 - F_{m,n}(x + h(x)) \le 1 - \Phi(x)$ i.e. the left-hand side of formula (2.1) holds.

Let $K = \frac{100K_1}{C_1}$. If we choose the constants $\bar{A} > 0$ and $B > 0$ (depending on the number K) sufficiently small, then the inequalities $|x + h_{m,n}(x)| < A\sqrt{n}$ and $|h(x)| \le x + 1$ hold. In this case

$$\begin{aligned} &\frac{K_1}{\sqrt{n}} \left((x + h(x))^3 + \frac{|n-m|}{\sqrt{n}} \left((x + h(x))^2 + x + h(x) \right) + 1 \right) \\ &\qquad \le \frac{K_1}{\sqrt{n}} \left((2x+1)^3 + \frac{|n-m|}{\sqrt{n}} \left((2x+1)^2 + 2x + 1 \right) + 1 \right) \\ &\qquad \le \frac{100K_1}{\sqrt{n}} \left(x^3 + \frac{|n-m|}{\sqrt{n}}(x^2+1) + 1 \right) \end{aligned}$$

$$\leq \frac{100K_1}{\sqrt{n}}(x+2)\left(x^2+\frac{|n-m|}{\sqrt{n}}(x+1)+1\right)=\frac{100K_1}{K}(x+2)h(x)$$
$$\leq C_1(x+2)h(x).$$

This inequality implies the left-hand side of formula (2.1). The right-hand side of this inequality can be proved similarly. Formula (2.1) implies that

$$F_{m,n}(x-h(x))\leq\Phi(x)\leq F_{m,n}(x+h(x)),\quad \text{if } 0\leq x\leq \bar A\sqrt{n}, \text{ and } |n-m|<Bn.$$

This inequality also holds if the condition $0\leq x\leq \bar A\sqrt{n}$ is replaced by the condition $0\geq x\geq -\bar A\sqrt{n}$, and it can be proved similarly. Thus we get that

$$x-h_{m,n}(x)\leq F^{-1}_{n,m}(\Phi(x))\leq x+h_{m,n}(x),\quad \text{if } |x|\leq \bar A\sqrt{n} \text{ and } |n-m|\leq Bn,$$

and

$$|F^{-1}_{m,n}(\Phi((\eta))-\eta|\leq h_{m,n}(\eta)=K\frac{\eta^2+\frac{|n-m|}{\sqrt{n}}(|\eta|+1)+1}{\sqrt{n}}\leq \bar K\frac{\eta^2+\frac{(n-m)^2}{n}+1}{\sqrt{n}},$$

with an appropriate constant $\bar K>0$ if $|\eta|\leq \bar A\sqrt{n}$ and $|n-m|\leq Bn$. (In the last estimation we have exploited that $\frac{|n-m|}{\sqrt{n}}(|\eta|+1)\leq\frac12\left(\frac{(n-m)^2}{n}+\eta^2+1\right)$.) In such a way he have solved problem 2 (with the notation $\bar A$ instead of A in the last relation).

Remark: Actually we have proved the following slightly stronger estimate:

$$|F^{-1}_{m,n}(\Phi(\eta))-\eta|\leq K\frac{\eta^2+\frac{|n-m|}{\sqrt{n}}(|\eta|+1)+1}{\sqrt{n}}\quad \text{on the set } \{|\eta|<A\sqrt{n}\}.$$

But the estimate formulated in problem 2 will be more convenient for us.

3.) The random variables $V_{k,l}(n)$, $k=1,\dots,2^l$, generate the same σ-algebra as the random variables $m_k=\frac{\sqrt{n}}{2^{(l+1)/2}}V_{k,l}(n)+\frac{n}{2^l}$, $k=1,\dots,2^l$. Let us consider that sequence $\zeta_1,\dots,\zeta_n$ of independent and on the interval $[0,1]$ uniformly distributed random variables which defines the standardized empirical function $Z_n(t)$, $0\leq t\leq 1$. Then the value of the random variable m_k equals the number of those elements of the sequence $\zeta_1,\dots,\zeta_n$ which fall into the interval $[(k-1)2^{-l},k2^{-l}]$. This relation holds, since $m_k=n[P_n(k2^{-l})-P_n((k-1)2^{-l})]$, where $P_n(t)$ is the empirical distribution function defined at the beginning of this series of problems.

Let us fix the values of the random variables m_k, $1\leq k\leq 2^l$, i.e. the atom $B(m_1,\dots,m_{2^l})$ of the σ-algebra $\mathcal{G}_l(n)$. The random variables of the values Y_k with respect to the σ-algebra $\mathcal{G}_l(n)$ and the conditional distribution of the random variable Y_k on the atom $B(m_1,\dots,m_{2^l})$ is the binomial distribution $B(m_k,\frac12)$ with parameters m_k and $\frac12$. This implies that

$$\begin{aligned}
V_{2k-1,l+1}(n) &= \frac{2^{(l+2)/2}}{\sqrt{n}}Y_k - \frac{\sqrt{n}}{2^{l/2}},\\
E\left(V_{2k-1,l+1}(n)|\mathcal{G}_l(n)\right) &= \frac{2^{(l+2)/2}}{\sqrt{n}}E(Y_k|\mathcal{G}_l(n)) - \frac{\sqrt{n}}{2^{l/2}} = \frac{2^{(l+2)/2}}{\sqrt{n}}\frac{m_k}{2} - \frac{\sqrt{n}}{2^{l/2}}\\
&= \frac{1}{\sqrt{2}}V_{k,l}(n)
\end{aligned}$$

and

$$\bar V_{2k-1,l+1}(n) = \frac{2^{(l+2)/2}}{\sqrt{n}}Y_k - \frac{2^{(l+2)/2}}{\sqrt{n}}\frac{m_k}{2}$$

on the atom $B(m_1,\dots,m_{2^l})$ of the σ-algebra $G_l(n)$. Furthermore,

$$\bar V_{2k-1,l+1}(n) + \bar V_{2k,l+1}(n) = V_{k,l}(n) - E(V_{k,l}(n)|\mathcal{G}_l(n)) = V_{k,l}(n) - V_{k,l}(n) = 0,$$

and

$$\begin{aligned}
E\left(V_{2k,l+1}(n)|\mathcal{G}_l(n)\right) &= \sqrt{2}E\left(V_{k,l}(n)|\mathcal{G}_l(n)\right) - E\left(V_{2k-1,l+1}(n)|\mathcal{G}_l(n)\right)\\
&= \left(\sqrt{2} - \frac{1}{\sqrt{2}}\right)V_{k,l}(n) = \frac{1}{\sqrt{2}}V_{k,l}(n).
\end{aligned}$$

The above relations imply the statement of the problem (with the choice $Y_k = X_{2k-1}$).

4.) As in this problem normally distributed random variables are considered, its statements can be proved by the investigation of the covariance function. The calculation needed for the proof can be simplified by using the representation $B(t) = W(t) - tW(1)$ of a Brownian bridge, where $W(t)$, $0 \le t \le 1$, $EW(t) = 0$, $EW(s)W(t) = \min(s,t)$ is a Wiener process. Simple calculation shows that

$$E\left(U_{2k-1,l+1} - \frac{1}{\sqrt{2}}U_{k,l}\right)U_{j,l} = E\left(U_{2k,l+1} - \frac{1}{\sqrt{2}}U_{k,l}\right)U_{j,l} = 0.$$

From this relation and the joint Gaussian distribution of the random variables under consideration imply that the random vector

$$\left\{U_{2k-1,l+1} - \frac{1}{\sqrt{2}}U_{k,l}, U_{2k,l+1} - \frac{1}{\sqrt{2}}U_{k,l},\ k = 1,\dots,2^l\right\}$$

is independent of the σ-algebra $\mathcal{F}_l$, and its coordinates are Gaussian random variables with expectation zero.

Hence $E\left(U_{2k-1,l+1}-\dfrac{1}{\sqrt{2}}U_{k,l}\middle|\mathcal{F}_l\right)=0$, and $E(U_{2k-1,l+1}|\mathcal{F}_l)=\dfrac{1}{\sqrt{2}}E(U_{k,l}|\mathcal{F}_l)=\dfrac{1}{\sqrt{2}}U_{k,l}$. Similarly, $E(U_{2k,l+1}|\mathcal{F}_l)=\dfrac{1}{\sqrt{2}}U_{k,l}$, $1\le k\le 2^l$. Hence, $\bar U_{2k-1,l+1}=U_{2k-1,l+1}-\dfrac{1}{\sqrt{2}}U_{k,l}$, $\bar U_{2k,l+1}=U_{2k,l+1}-\dfrac{1}{\sqrt{2}}U_{k,l}$, $k=1,\dots,2^l$.

Simple calculation shows that $\bar U_{2k-1,l+1}=\frac{1}{\sqrt{2}}(U_{2k-1,l+1}-U_{2k,l+1})=-\bar U_{2k,l+1}$ and $E\bar U_{2j-1,l+1}\bar U_{2k-1,l+1}=\delta_{j,k}$, $1\le k\le 2^l$, where $\delta_{j,k}=0$ if $j\ne k$ and $\delta_{j,k}=1$ if $j=k$. This means that the random variables $\bar U_{2k-1,l+1}$, $k=1,\dots,2^l$, are independent of the σ-algebra $\mathcal{F}_l$ and they have standard normal distribution. Beside this, the identity $\bar U_{2k-1,l+1}=-\bar U_{2k,l+1}$ holds. In such a way we have solved problem 4.

5a.) We prove the statement of problem 5a by induction with respect to the parameter L. For $L=0$ the statements of problem 5a hold. Let us assume that we have already proved them for $L=l$. Then we want to prove them for $L=l+1$. First we show that the random variables $Z_n(k2^{-(l+1)})$, $1\le k\le 2^{l+1}$, we construct have the right joint distributions. The proof we give may be a little complicated, but there is a simple idea behind it. We compare the random variables we have to handle with analogous random variables constructed by means of a standardized empirical distribution functions. Then we check that the definition of the random variables given in formulas (5a)–(5f) guarantee enough similarities to prove the desired results.

Let us define the random variables $M_k=\dfrac{\sqrt{n}}{2^{(l+1)/2}}V_{k,l}(n)+\dfrac{n}{2^l}$, $k=1,\dots,2^l$. The random variables $\bar V_{2k-1,l+1}(n)$ defined in formula (5b) are transforms of the random variables $\bar U_{2k-1,l+1}$ which are by the results of problem 4 independent random variables with standard normal distribution. Furthermore, they are also independent of the σ-algebra $\mathcal{G}_l(n)$, since they are independent of the σ-algebra $\mathcal{F}_l\supset\mathcal{G}_l(n)$. These facts imply that the random variables $\bar V_{2k-1,l+1}$ are conditionally independent with respect to the random variables M_k, $1\le k\le 2^l$, which generate the σ-algebra $\mathcal{G}_l(n)$. Beside this, the conditional distribution of the random variables $\bar V_{2k-1,l+1}$ with respect to the condition $M_k=m_k$, $1\le k\le 2^l$, is the distribution functions $F_{m_k,l}(x)$ defined in formulas (5a) and (5a$'$).

Let us fix a standardized empirical distribution function $Z_n(t)$, and consider the random variables $V_{\cdot,\cdot}$, $\bar V_{\cdot,\cdot}$ and σ-algebra $\mathcal{G}_\cdot$ defined from this process by formulas (1)—(4), where the subscript marks "$\cdot$" in these formulas denote some indices. Let us also define the random variables $\bar M_k$ similarly to the random variables M_k defined in the previous paragraph with the difference that now we replace the random variable $V_{k,l}(n)$ considered there by the random variable $V_{k,l}(n)$ determined by that standardized empirical process $Z_n(t)$ which is considered in this paragraph. Our goal is to show by a comparison of the random variables defined in this paragraph from a standardized empirical function with the random variables defined in problem 5a that the latter variables have the right distributions. To do this let us first observe that because of the results of problem 3 the random vector $\bar V_{2k-1,l+1}(n)$, $1\le k\le 2^l$, defined in this paragraph has the same conditional distribution with respect to the conditions $\bar{M}_k = m_k$, $1 \le k \le 2^l$, i.e. on the atom $B(m_1, \dots, m_{2^l})$ of the σ-algebra $\mathcal{G}_l(n)$ as the random variables $\bar{V}_{2k-1,l}(n)$ constructed in formula (5b) with respect to the conditions $M_k = m_k$, $1 \le k \le 2^l$. Furthermore, the random variables considered in this paragraph satisfy the relation $E(V_{2k-1,l+1}(n)|\mathcal{G}_l(n)) = \frac{1}{\sqrt{2}} V_{k,l}(n)$ by the results of problem 3. Hence a comparison of formulas (1)—(4) with formulas (5c) and (5d) implies the following statement: Take the conditional distribution of the random vector $Z_n((2k-1)2^{-(l+1)})$, $1 \le k \le 2^l$, considered in problem 5a under the condition that the previously constructed random variables $Z_n(k2^{-l})$, $1 \le k \le 2^l$ have prescribed values. This conditional distribution agrees with the analogous conditional distribution which we get by replacing the values of the random variables $Z_n((2k-1)2^{-(l+1)})$ and $Z_n(k2^{-l})$, $1 \le k \le 2^l$, by the values of a standardized empirical process $Z_n(t)$, $0 \le t \le 1$, in the corresponding points.

Let us make the following observation. The joint distribution of the random variables $Z_n(k2^{-l})$, $1 \le k \le 2^l$, together with the conditional distribution of the random vector $Z_n((2k-1)2^{-(l+1)})$, $1 \le k \le 2^l$, with respect to the condition that the values of the random variables $Z_n(k2^{-l})$, $1 \le k \le 2^l$, take their prescribed values determine the joint distribution of the random variables $Z_n(k2^{-(l+1)})$, $1 \le k \le 2^{l+1}$. Hence the results about the conditional distribution of the random vector $Z_n((2k-1)2^{-(l+1)})$, $1 \le k \le 2^l$, and the induction hypothesis imply that the distribution of the random vector $Z_n(k2^{-(l+1)})$, $1 \le k \le 2^{l+1}$, constructed in this problem 5a agrees with the joint distribution of a standardized empirical distribution function $Z_n(t)$ in the points $t = k2^{-(l+1)}$, $1 \le k \le 2^{+1}$.

Now we show that the above defined random variables $Z_n(k2^{-(l+1)})$, $V_{k,l+1}(n)$ and $\bar{V}_{k,l+1}(n)$ satisfy formulas (1)—(4). The random variables defined in formula (5b) satisfy the identity $E(\bar{V}_{2k-1,l}(n)|\mathcal{G}_l(n)) = 0$. To see this let us consider the conditional distribution of the random variable $\bar{V}_{2k-1,l}(n)$ with respect to the σ-algebra $\mathcal{G}_l(n)$ on that atom, where

$$m_k = \frac{\sqrt{n}}{2^{(l+1)/2}} V_{k,l}(n) + \frac{n}{2^l}, \quad k = 1, \dots, 2^l.$$

This conditional distribution function is the distribution function $F_{m_k,l}(x)$ defined in (5b), hence the conditional expectation of a random variable with such a conditional distribution equals zero. This relation together with formula (5c) imply that $E(V_{2k-1,l+1}(n)|\mathcal{G}_l(n)) = \frac{1}{\sqrt{2}} V_{k,l}(n)$, and

$$\bar{V}_{2k-1,l+1}(n) = V_{2k-1,l+1}(n) - E(V_{2k-1,l+1}(n)|\mathcal{G}_l(n)).$$

By formula (5d)

$$V_{2k-1,l+1}(n) = 2^{(l+2)/2} \left(Z_n\left(\frac{2k-1}{2^{l+1}}\right) - Z_n\left(\frac{2k-2}{2^{l+1}}\right) \right).$$

The above identities contain the statements of formulas (1)—(4) to be proved for $L = l+1$ if only the odd indices k are considered. To prove the corresponding identities for even indices k let us first observe that it follows from the last identity, formula (5e) and formula (1b) already proved for index l that

$$\begin{aligned} V_{2k,l+1}(n) &= 2^{(l+2)/2}\left(Z_n\left(\frac{k}{2^l}\right)-Z_n\left(\frac{k-1}{2^l}\right)\right) \\ &\qquad - 2^{(l+2)/2}\left(Z_n\left(\frac{2k-1}{2^{l+1}}\right)-Z_n\left(\frac{2k-2}{2^{l+1}}\right)\right) \\ &= 2^{(l+2)/2}\left(Z_n\left(\frac{2k}{2^{l+1}}\right)-Z_n\left(\frac{2k-1}{2^{l+1}}\right)\right). \end{aligned}$$

By applying again formula (5e), and the relation already proved for the random variable $V_{2k-1,l+1}(n)$ we get that

$$\begin{aligned} E(V_{2k,l+1}(n)|\mathcal{G}_l(n)) &= \sqrt{2}V_{k,l}(n) - E(V_{2k-1,l+1}(n)|\mathcal{G}_l(n)) \\ &= \sqrt{2}V_{k,l}(n) - \frac{1}{\sqrt{2}}V_{k,l}(n) = \frac{1}{\sqrt{2}}V_{k,l}(n), \end{aligned}$$

and

$$V_{2k,l+1}(n) = \sqrt{2}V_{k,l}(n) - \bar V_{2k-1,l+1}(n) - \frac{1}{\sqrt{2}}V_{k,l}(n) = \frac{1}{\sqrt{2}}V_{k,l}(n) - \bar V_{2k-1,l+1}(n).$$

It follows from these formulas and relation (5f) that also the identity $\bar V_{2k,l+1}(n) = -\bar V_{2k-1,l+1}(n) = V_{2k,l+1}(n) - E(V_{2k,l+1}(n)|\mathcal{G}_l(n))$ holds. In such a way we have proved formulas (1)—(4) also in $l+1$-th step.

Let us finally observe that also the relation $\mathcal{G}_{l+1}(n) \subset \mathcal{F}_{l+1}$ holds, since the random variables $V_{k,l+1}(n)$ generating the σ-algebra $\mathcal{G}_{l+1}(n)$ are measurable functions of the $\mathcal{F}_{l+1}$ measurable random variables $U_{k,l+1}$.

5b.) It follows immediately from the construction of the random variables ζ_j, $1\le j\le n$, given in problem 5b that the values of the standardized empirical distribution function made from it agree with the random variables $Z_n(k2^{-L})$ constructed in problems 5a and 5b for all numbers $1\le k\le 2^L$. We still have to show that they constitute a sequence of independent random variables with uniform distribution in the interval $[0,1]$.

To show this first we prove the following statement. If $\bar\zeta_1,\dots,\bar\zeta_n$ is a sequence of independent and in the interval $[0,1]$ uniformly distributed random variables, $\bar\zeta_1^*\le\cdots\le\bar\zeta_n^*$ is the ordered sample made from this random variables, and M_k denotes the number of those points of the sequence $\bar\zeta_j^*$, $1\le j\le n$, which fall into the interval $[(k-1)2^{-L},k2^{-L}]$, $1\le k\le 2^L$, then $P(M_1=m_1,\dots,M_{2^L}=m_{2^L}) = \dfrac{n!}{m_1!\cdots m_{2^L}!}2^{-Ln}$, if $m_k\ge 0$ for all indices $1\le k\le 2^L$, and $\sum\limits_{k=1}^{2^L} m_k=n$. Furthermore, the conditional distribution of the random sequence $\bar\zeta_1^*\le\cdots\le\bar\zeta_n^*$

under the condition that $\{M_1 = m_1, \dots, M_{2^L} = m_{2^L}\}$ agrees with the distribution of the union of such independent sequences of random variables

$$\xi_{m_0+\cdots+m_{k-1}+1}, \xi_{m_0+\cdots+m_{k-1}+2}, \dots, \xi_{m_0+\cdots+m_k}, \quad 1 \le k \le 2^L,\ m_0 = 0,$$

independent of each other, for which

$$\xi_{m_0+\cdots+m_{k-1}+1}, \xi_{m_0+\cdots+m_{k-1}+2}, \dots, \xi_{m_0+\cdots+m_k}$$

is an ordered sample made from m_k on the interval $[(k-1)2^{-l}, k2^{-L}]$ uniformly distributed and independent random variables. Indeed, it is easy to check the identity

$$P(M_1 = m_1, \dots, M_{2^L} = m_{2^L}) = \frac{n!}{m_1! \cdots m_{2^L}!} 2^{-Ln}$$

On the other hand, if we prescribe for all random variables $\bar\zeta_j$, $1 \le j \le n$, which interval $[m(j)2^{-L}, (m(j)+1)2^{-L}]$ they fall into, and we do this in such a way that the number of points $\bar\zeta_j$ falling into the interval $[(k-1)2^{-L}, k2^{-L}]$ equals m_k, $1 \le k \le 2^L$, then the random variables $\bar\zeta_j$, $1 \le j \le n$, are conditionally independent under this condition, and $\bar\zeta_j$ is uniformly distributed in the interval $[m(j)2^{-L}, (m(j)+1)2^{-L}]$. It follows from this relation that the conditional distribution of the ordered sample a $\bar\zeta_1^* \le \cdots \le \bar\zeta_n^*$ under such a condition, hence also under the union of such conditions, i.e. under the condition $M_1 = m_1, \dots, M_{2^L} = m_{2^L}$ equals the above described conditional distribution.

Let us consider the random sequence $\zeta_1^* \le \cdots \le \zeta_n^*$ constructed in problem 5b and define the random variables M'_k, where M'_k equals the number of points of the sequence $[(k-1)2^{-L}, k2^{-L}]$ which fall into the interval $[(k-1)2^{-L}, k2^{-L}]$. Then $P(M'_1 = m_1, \dots, M'_{2^L} = m_{2^L}) = \dfrac{n!}{m_1! \cdots m_{2^L}!} 2^{-Ln}$, if $m_k \ge 0$ for all indices $1 \le k \le 2^L$, and $\sum\limits_{k=1}^{2^L} m_k = n$. Furthermore, the conditional distribution of the random sequence $\zeta_1^* \le \cdots \le \zeta_n^*$ under the condition $\{M'_1 = m_1, \dots, M'_{2^L} = m_{2^L}\}$ agrees with the distribution of the union of such sequences

$$\xi_{m_0+\cdots+m_{k-1}+1}, \xi_{m_0+\cdots+m_{k-1}+2}, \dots, \xi_{m_0+\cdots+m_k}, \quad 1 \le k \le 2^L,\ m_0 = 0,$$

which are independent of each and

$$\xi_{m_0+\cdots+m_{k-1}+1}, \xi_{m_0+\cdots+m_{k-1}+2}, \dots, \xi_{m_0+\cdots+m_k}$$

is an ordered sample of m_k independent and in the interval $[(k-1)2^{-l}, k2^{-L}]$ uniformly distributed random variables. Since the distribution of the sequence $M'_1, \dots, M'_{2^L}$ together with the conditional distribution of the random sequence $\zeta_1^* \le \cdots \le \zeta_n^*$ with respect to the condition $\{M'_1 = m_1, \dots, M'_{2^L} = m_{2^L}\}$ determines the joint distribution of the random sequence $\zeta_1^* \le \cdots \le \zeta_n^*$ the previous relations imply that the joint distribution of the sequences $\bar\zeta_1^* \le \cdots \le \bar\zeta_n^*$ and $\zeta_1^* \le \cdots \le \zeta_n^*$ agree. This means that $\zeta_1^* \le \cdots \le \zeta_n^*$ is the ordered sample made from n independent and on the interval $[0,1]$ uniformly distributed random variables.

Let us finally make the following observation. Let $\bar\zeta_1,\dots,\bar\zeta_n$ be a sequence of independent and in the interval $[0,1]$ uniformly distributed random variables, and define a random permutation of the set $\{1,\dots,n\}$ by means of this sequence though the following relation. $\bar\zeta_j=\bar\zeta^*_{\pi(j)}$, $j=1,\dots,n$, where $\bar\zeta_1^* \le \cdots \le \bar\zeta_n^*$ is the ordered sample $\bar\zeta_1^* \le \cdots \le \bar\zeta_n^*$ made from this sequence. Then the random vectors $(\pi(1),\dots,\pi(n))$, and $(\bar\zeta_1^*,\dots,\bar\zeta_n^*)$ are independent, and the vector $(\pi(1),\dots,\pi(n))$ is uniformly distributed on the permutations of the set $\{1,\dots,n\}$. From this fact and the observation that we have to handle an ordered sample $\zeta_1^* \le \cdots \le \zeta_n^*$ made from independent and in the interval $[0,1]$ uniformly distributed random variables together with a uniformly distributed permutation $(\pi(1),\dots,\pi(n))$ follows that the sequence $(\zeta_1,\dots,\zeta_n)=(\zeta^*_{\pi(1)},\dots,\zeta^*_{\pi(n)})$ constructed in problem 5b consists of independent and in the interval $[0,1]$ uniformly distributed random variables.

6.) The main statement of problem 6 is a consequence of the result of problem 2. In order not to denote different quantities with the same letter we apply the result of problem 2 with the notation $\bar m$ and $\bar n$ instead of the letters m and n.

Let us apply the result of the second problem with the choice $\bar m=\frac{n}{2^l}$, $\bar n=m_k=\frac{\sqrt n}{2^{(l+1)/2}}V_{k,l}(n)+\frac{n}{2^l}$ and $\eta=\bar U_{2k-1,l+1}$. Then $|\bar n-\bar m|=\frac{\sqrt n}{2^{(l+1)/2}}|V_{k,l}(n)|=\frac{\sqrt{\bar m}}{\sqrt2}|V_{k,l}(n)|$. Simple calculation shows that the conditions $|\bar n-\bar m|<B\bar n$ and $|\eta|<A\sqrt{\bar n}$ of problem 2 hold under the conditions of problem 6 if the constant $A>0$ appearing at the end of its formulation is chosen sufficiently small. On the other hand, $\bar V_{2k-1,l+1}(n)-\bar U_{2k-1,l+1}=F^{-1}_{m_k,l}(\Phi(\eta))-\eta$ by formula (5b) with the function $F_{m_k,l}(x)$ defined in (5a′) which agrees with the function $F_{\bar n,\bar m}(x)$ in problem 2. Beside this the appropriate term in the estimate of problem 2 satisfies the relation $\frac{(\bar n-\bar m)^2}{\bar n}=\frac{V_{k,l}(n)^2}{2}\frac{\bar m}{\bar n}=\text{const.}\,V_{k,l}(n)^2$. Hence the first part of relation (7a) follows from the result of problem 2. Its second part is a simple consequence of the relation $\bar U_{2k-1,l+1}=-\bar U_{2k-1,l+1}$. The inequality (7b) is a consequence of formula (7a), since

$$|U_{2k-1,l+1}-V_{2k-1,l+1}(n)|\le|\bar U_{2k-1,l+1}-\bar V_{2k-1,l+1}(n)|+\frac{|U_{k,l}-V_{k,l}(n)|}{\sqrt2},$$

and

$$|U_{2k,l+1}-V_{2k,l+1}(n)|\le|\bar U_{2k-1,l+1}-\bar V_{2k-1,l+1}(n)|+\frac{|U_{k,l}-V_{k,l}(n)|}{\sqrt2}.$$

These relations hold, since $V_{2k-1,l}(n)=\bar V_{2k-1,l}(n)+\frac{V_{k,l}(n)}{\sqrt2}$, $U_{2k-1,l}=\bar U_{2k-1,l}+\frac{U_{k,l}}{\sqrt2}$ by the identities $E(V_{2k-1,l}(n)|\mathcal G_l(n))=\frac{V_{k,l}(n)}{\sqrt2}$ and $E(U_{2k-1,l}(n)|\mathcal F_l)=\frac{U_{k,l}}{\sqrt2}$

proved in problems 3 and 4. Beside this a similar relation holds also for the random variables $V_{2k,l}(n)$ and $U_{2k,l}$.

7.) By the definition of the random variables $V_{k,l}$ and numbers $\varepsilon(j)$ and $k(j)$ the identity

$$\varepsilon(j)2^{-(j+1)/2}\left(\sqrt{2}V_{k_{j-1}+1,j-1}(n)-V_{k_j+1,j}(n)\right) \\ =\varepsilon(j)\left(Z_n(k_j2^{-j})-Z_n(k_{j-1}2^{-(j-1)})\right)$$

holds for all numbers $1\le j\le l$. Indeed, either $\varepsilon(j)=0$ when the above identity is obvious or $\varepsilon(j)=1$ when $k_j=2k_{j-1}+1$ and the identity follows from the definition of the random variables $V_{k,l}(n)$. By summing up these identities and exploiting the relations $Z_n(k_l2^{-l})=Z_n(t)$ and $Z_n(k_0)=Z_n(0)=0$ we get the first line of formula (8a) about the representation of the random variable $Z_n(t)$. The analogous formula about the expression $B(t)$ can be proved similarly.

In the proof of relation (8b) we apply formula (7b) with the choice $l=j-s-1$ and $k=k_{j-s-1}+1$. If $u_{j-s-1}=k_{j-s-1}2^{-j-s-1}=u_{j-s}=k_{j-s}2^{-j-s}$, then $k_{j-s}+1=2(k_{j-s-1}+1)-1$, and we consider the first term at left-hand side term together with the first expression at the right-hand side of this inequality. If $u_{j-s-1}=k_{j-s-1}2^{-j-s-1}<u_{j-s}=k_{j-s}2^{-j-s}$, then $k_{j-s}+1=2(k_{j-s-1}+1)$. In this case we consider the second term at the left-hand side together with the second expression at the right-hand side of this inequality. With such a choice we get that

$$\begin{aligned}&2^{-s}\cdot 2^{-(j-s+1)/2}|U_{k_{j-s}+1,j-s}-V_{k_{j-s}+1,j-s}(n)|\\ &\quad<2^{-s}\cdot\frac{K}{\sqrt{n}}(\bar U^2_{k_{j-s}+1,j-s}+V^2_{k_{j-s-1}+1,j-s-1}(n)+1)\\ &\quad\quad+2^{-(s+1)}\cdot 2\cdot\frac{2^{-(j-(s+1)+1)/2}}{2}|U_{k_{j-(s+1)}+1,j-(s+1)}-V_{k_{j-(s+1)}+1,j-(s+1)}(n)|\end{aligned}$$

for all pairs $1\le j\le l$ and $0\le s\le j-1$. We get inequality (8b) by summing up these inequalities for all numbers $0\le s\le j-1$. To see this we have to observe that after this summation for all indices $1\le j-s\le j-1$ the terms $|U_{k_{j-s}+1,j-s}-V_{k_{j-s}+1,j-s}(n)|$ appear with the same coefficient on the two sides of these sums. Beside this, we have to check that if $\omega\in\mathbf{B}$, where $\mathbf{B}$ is the set defined in problem 7, then the previous inequalities hold, i.e. the random variables $\bar U_{k_{j-s}+1,j-s}(\omega)$ and $V_{kj-s+1,j-s}(\omega)$ satisfy the conditions of problem 6.

The random variable $Z_n(k2^{-l})-B(k2^{-l})$ can be expressed by means of formula (8a) formula as a linear combination of the expressions $V_{k_j+1,j}(n)-U_{k_j+1,j}$. All these terms can be bounded by means of formula (8b). By applying these estimations we get the formula

$$|Z_n(k2^{-l})-B(k2^{-l})|\le 2\sum_{j=1}^{l}2^{-(j+1)/2}|V_{k_j+1,j}(n)-U_{k_j+1,j}|$$

$$\leq \frac{2K}{\sqrt{n}} \sum_{j=1}^{l} \sum_{s=0}^{j-1} 2^{-s} \left(\bar{U}^2_{k_{j-s}+1,j-s} + V^2_{k_{j-s-1}+1,j-s-1}(n) + 1 \right)$$
$$= \frac{2K}{\sqrt{n}} \sum_{j=1}^{l} \sum_{s=1}^{j} 2^{-(j-s)} \left(\bar{U}^2_{k_s+1,s} + V^2_{k_{s-1}+1,s-1}(n) + 1 \right)$$
$$= \frac{2K}{\sqrt{n}} \sum_{s=1}^{l} \left(\bar{U}^2_{k_s+1,s} + V^2_{k_{s-1}+1,s-1}(n) + 1 \right) \sum_{j=s}^{l} 2^{-(j-s)}.$$

Formula (8c) follows from this relation.

8.) By the results of problem 4 $\bar{U}_{k,j}$, $1 \leq j \leq l$, $1 \leq k \leq 2^j$, are independent random variables with standard normal distribution. This implies the statement of the problem about the joint distribution of the random variables $\bar{U}_{k_j+1,j}$, $1 \leq j \leq l$. The analogous statement about the joint distribution of the random variables $V_{k_{j-1}+1,j-1}(n)$, $1 \leq j \leq l$, follows from the statement formulated below by means of induction with respect to the parameter j.

We claim that the conditional distribution of the random variable $V_{k_{j-1}+1,j-1}(n)$ with respect to the conditions $M_s = \frac{\sqrt{n}}{2^{(s+1)/2}} V_{k_s+1,s} - \frac{n}{2^s} = m_s$, $1 \leq s \leq j-2$, agrees with the conditional distribution of the random variable $V_{1,j-1}(n)$, with respect to the conditions $\bar{M}_s = \frac{\sqrt{n}}{2^{(s+1)/2}} V_{1,s} - \frac{n}{2^s} = m_s$, $1 \leq s \leq j-2$, where $m_1, \dots, m_{s-2}$, are arbitrary non-negative integers. This statement can be rewritten in an equivalent form by expressing the conditional distributions of the random variables $M_{j-1} = \frac{\sqrt{n}}{2^{j/2}} V_{k_{j-1}+1} - \frac{n}{2^{j-1}}$ and $\bar{M}_{j-1} = \frac{\sqrt{n}}{2^{j/2}} V_{1,j-1} - \frac{n}{2^{(j-1)}}$ with respect to the conditions $M_s = m_s$ and $\bar{M}_s = m_s$, $1 \leq s \leq j-1$ respectively instead of working with the random variables $V_{k_{j-1}+1,j-1}(n)$ and $V_{1,j-1}(n)$. This modified statement follows from the observation that the conditional distribution both of M_{j-1} and $\bar{M}_{j-1}$ with respect to the appropriate conditions is the binomial distribution $B(m_{j-2}, \frac{1}{2})$ with parameters m_{j-2} and $\frac{1}{2}$.

The inequality $1 - P(\mathbf{B}) \leq e^{-D_1 x}$ will be proved by means of the following estimations. For arbitrary numbers $1 \leq j \leq l$

$$P\left(|V_{k_{j-1}+1,j-1}(n)| > \frac{A\sqrt{n}}{2^{j/2}} \right) = P\left(\frac{2^{j/2}}{\sqrt{n}} \left| \sum_{k=1}^{n} (\chi_k - E\chi_k) \right| > \frac{A\sqrt{n}}{2^{j/2}} \right)$$
$$= P\left(\left| \sum_{k=1}^{n} (\chi_k - E\chi_k) \right| > \frac{An}{2^j} \right),$$

where χ_k, $1 \leq k \leq n$, are independent random variables, and $P(\chi_k = 1) = 1 - P(\chi_k = 0) = 2^{-(j-1)}$. This relation implies that

$$Ee^{t(\chi_k - E\chi_k)} = \left(1 - \frac{1}{2^{j-1}} + \frac{e^t}{2^{j-1}} \right) e^{-t/2^{j-1}},$$

and from here $Ee^{t(\chi_k - E\chi_k)} \le \exp\left\{\frac{e^t-1}{2^{j-1}} - \frac{t}{2^{j-1}}\right\} \le \exp\left\{\frac{10t^2}{2^j}\right\}$ if $|t|<1$. In the latter estimations we exploited that the inequalities $t+1\le e^t$ and $e^t-1-t<5t^2$ hold if $|t|\le 1$. By the previous estimates $E\left(\exp\left\{t\sum\limits_{k=1}^n(\chi_k - E\chi_k)\right|\right\}\le e^{10t^2n2^{-j}}$ if $|t|<1$, and

$$P\left(V_{k_{j-1}+1,j-1}(n) > \frac{A\sqrt{n}}{2^{j/2}}\right) = P\left(\exp\left\{t\sum_{k=1}^n(\chi_k - E\chi_k)\right\} > e^{Ant/2^j}\right) \tag{2.2a}$$
$$\le e^{n2^{-j}(10t^2-At)} \le e^{-\bar D 2^{l-j}x} \quad \text{for all numbers } 1\le j\le l.$$

with an appropriate constant $\bar D>0$ if $j\le l$. In the last inequality we have exploited that $10t^2-At<-D'$ with an appropriate constant $D'>0$, if the number $t>0$ is chosen sufficiently small, and $-n2^{-j}=-2^{l-j}2^{-l}n\le -Cx2^{l-j}$ under the conditions of problem 8.

As $\bar U_{k_{j-1}+1,j-1}$ is a random variable with standard normal distribution, hence simple calculation yields that

$$P\left(|U_{k_{j-1}+1,j-1}(\omega)| < \frac{A\sqrt{n}}{2^{j/2}}\right) \le e^{-A^2n2^{j-1}} \le e^{-D''2^{l-j}x} \tag{2.2b}$$
$$\text{for all numbers } 1\le j\le l.$$

The estimate (2.2a) remains valid if the random variable $V_{k_{j-1}+1,j-1}(n)$ is replaced by $-V_{k_{j-1}+1,j-1}(n)$. If we sum up the inequalities (2.2a) theirs analogs formulated above and the inequality (2.2b) for all numbers $1\le j\le n$, then we get the inequalities $1-P(\mathbf{B})\le e^{-D_1x}$.

To prove the last estimate of problem 8 let us consider the square of a normally distributed random variable, and let us calculate the moment generating function of the normalization of the random variables formulated in such a way. If η is a random variable with standard normal distribution, then

$$Ee^{t(\eta^2-E\eta^2)} = Ee^{t(\eta^2-1)} = \frac{e^{-t}}{\sqrt{2\pi}}\int e^{tx^2-x^2/2}\,dx = \frac{e^{-t}}{\sqrt{1-2t}}, \quad \text{if } t<\frac12.$$

It follows from here that $\log Ee^{t(\eta^2-E\eta^2)} = -t-\frac12\log(1-2t) < t^2$, if $0<t\le\frac14$. (The deeper reason for the validity of such an estimate is that the moment generating function of a random variable with expectation zero behaves in a small neighbourhood of the zero as $e^{\text{const.}\,t^2}$.) Hence

$$P\left(18K\sum_{j=1}^l \bar U_{1,j-1}^2 > x\right) = P\left(\exp\left\{\frac14\sum_{j=1}^l(\bar U_{1,j-1}^2 - E\bar U_{1,j-1}^2)\right\} > e^{x/72K}\right)$$
$$\le e^{l/16-x/72K} \le e^{-x/144K} = e^{-D_2x},$$

as we formulated in the problem, under the condition $l \le \frac{x}{9K}$. This inequality holds, since under the conditions of problem 8 $l \le \log C + \log n - \log x \le 2\log n \le \frac{2x}{C_0} \le \frac{x}{9K}$ if the constant C_0 is chosen sufficiently large in these conditions, and $n \ge n_0$ with an appropriate index n_0.

9.) Inequality (8c) holds for all $t = k2^{-l}$, $k = 1, \dots, 2^l$, on the set $\mathbf{B}_0 = \bigcap\limits_{k=1}^{2^l} \mathbf{B}(k2^{-l})$, where the set $\mathbf{B}(t)$ was defined after formula (8c). On the basis of the results of problem 8 the inequality $1 - P(\mathbf{B}_0) \le 2^l e^{-D_1 x} \le \frac{n}{Cx} e^{-D_1 x} \le n e^{-D_1 x} \le e^{x/C_0 - D_1 x} \le e^{-D_1 x/2}$ holds if the constant C_0 and the threshold index n_0 are chosen sufficiently large in the condition $x \ge C_0 \log n$ of problem 9.

Hence inequality (8c) and the first part of problem 8 imply that

$$
\begin{aligned}
P&\left(\sup_{1\le k\le 2^l} \sqrt{n}|Z_n(k(2^{-l}) - B(k2^{-l})| > \frac{x}{2}\right) \\
&\le e^{-D_1 x/2} + 2^l P\left(4K\sum_{j=1}^{l}\left(\bar U_{1,j}^2 + V_{1,j-1}^2(n) + 1\right) \ge \frac{x}{2}\right) \\
&\le e^{-D_1 x/2} + 2^l P\left(18K\sum_{j=1}^{l}\bar U_{1,j-1}^2 > x\right) + 2^l P\left(18K\sum_{j=1}^{l} V_{1,j-1}^2 > x\right).
\end{aligned}
$$

In the last inequality we exploited that $4K\sum\limits_{j=1}^{l} 1 = 4Kl \le \frac{x}{18}$, because $l \le \log n \le \frac{x}{C_0} \le \frac{x}{72K}$, if in the condition of relation (9) the constant $C_0 > 0$ and threshold index n_0 for which $n \ge n_0$ are chosen sufficiently large. Hence on the basis of formula (10), the last inequality of problem 8 and the previous estimate the left-hand side of the expression (9) can be bounded from above by the $e^{D_1 x/2} + 2^l\left(e^{-D_2 x} + e^{-D_3 x}\right)$. As $2^l \le n \le e^{\min(D_2, D_3)x/2}$, if the constant $C_0 > 0$ and threshold index n_0 are appropriately chosen in formula (9), this implies the statement of problem 9.

10.) As $Z_n(t) = X_n(t) - Y_n(t)$, $Z_n(t)^2 \le 2X_n(t)^2 + 2Y_n(t)^2$, and

$$
\begin{aligned}
P&\left(18\sum_{j=1}^{l} 2^j Z_n^2\left(\frac{1}{2^{j-1}}\right) > x\right) \\
&\le P\left(36\sum_{j=1}^{l} 2^j\left(X_n^2\left(\frac{1}{2^{j-1}}\right) + Y_n^2\left(\frac{1}{2^{j-1}}\right)\right) > x\right) \\
&\le P\left(72K\sum_{j=1}^{l} 2^j X_n\left(\frac{1}{2^{j-1}}\right)^2 > x\right) + P\left(72K\sum_{j=1}^{l} 2^j Y_n\left(\frac{1}{2^{j-1}}\right)^2 > x\right).
\end{aligned}
$$

This is formula (12).

If κ_n is a Poisson distributed random variable with parameter n, then the moment generating function of the random variable $\kappa_n - n$ is

$$Ee^{t(\kappa_n - n)} = e^{-tn} \sum_{k=0}^{\infty} \frac{n^k}{k!} e^{-n+tk} = e^{n(e^t - 1 - t)}.$$

This implies because of the inequality $n(e^t - t - 1) \le t^2$, under the condition $|t| \le 1$ that $Ee^{t(\kappa_n - n)} \le e^{nt^2}$, if $|t| \le 1$. Hence $P(\kappa_n - n > y) \le e^{nt^2 - ty} \le e^{-y^2/4n}$ with the choice $t = \frac{y}{2n}$, if $y \le 2n$. Similarly, $P(\kappa_n - n < -y) \le e^{-y^2/4n}$. The slightly more general statement formulated about $\kappa_n - n$ also holds. Indeed, if the condition $|y| \le 2n$ is replaced by the condition $|y| \le B_1 n$, then the previous relation remains valid if the exponent $-y^2/4n$ is replaced by the exponent $-B_2 y^2/n$ with an appropriate constant $B_2 > 0$.

To prove formula (13) let us observe that the random variables $Y_n(t)$ in (11c) are non-negative, hence

$$P\left(72K \sum_{j=1}^{l} 2^j Y_n\left(\frac{1}{2^{j-1}}\right)^2 > x\right) \le P\left(\left(\sqrt{72K} \sum_{j=1}^{l} 2^{j/2} Y_n\left(\frac{1}{2^{j-1}}\right)\right)^2 > x\right).$$

Then taking the conditional probability at the left-hand side of the next formula with respect to the condition $\kappa_n - n = m$, $-\infty < m < \infty$, and using the definition of the random variables $\bar{Y}_{n,m}$ introduced before formula (13) we get the following relation (2.3).

$$\begin{aligned} &P\left(\sqrt{72K} \sum_{j=1}^{l} 2^{j/2} \left| Y_n\left(\frac{1}{2^{j-1}}\right)\right| > \sqrt{x}\right) \\ &\quad = \sum_{m=-\infty}^{\infty} P\left(\sqrt{72K} \sum_{j=1}^{l} 2^{j/2} \bar{Y}_{n,|m|}\left(\frac{1}{2^{j-1}}\right) > \sqrt{x}\right) P(\kappa_n - n = m) \\ &\quad \le P\left(\sqrt{72K} \sum_{j=1}^{l} 2^{j/2} \bar{Y}_{n,B\sqrt{nx}}\left(\frac{1}{2^{j-1}}\right) > \sqrt{x}\right) P(|\kappa_n - n| \le B\sqrt{nx}) \\ &\qquad + P(|\kappa_n - n| > B\sqrt{nx}). \end{aligned} \tag{2.3}$$

The last relation of formula (2.3) can be seen with the help of the observation that the first probability in the second line of formula (2.3) is a monotone increasing function of the parameter $|m| = |\kappa_n - n|$. By fixing the number $M = B\sqrt{nx}$, and by replacing the parameter m by M in the case $|m| \le M$ and applying the trivial upper bound 1 for this probability we get the last inequality in formula (2.3).

The inequality in formula (13) is a simple consequence of relation (2.3) and the formula before it. Finally the identity formulated at the end of formula (13) follows from a simple calculation if we put the definition of the random variable $\bar{Y}_{n,B\sqrt{nx}}$ in the expression at the second line of formula (13), change the order of summation in the double sum obtained in such a way and put together the terms depending on the variable ζ_k, $1\le k\le B\sqrt{nx}$, in the form of a single random variable ξ_k.

The random variables $\xi_k=\xi_{k,l}$ are functions of the independent random variables ζ_k. It follows from this fact and the explicit form of the definition of the random variables ξ_k that they are independent and identically distributed.

11.) If the conditions of the problem are satisfied, then $2^{l/2}\le\sqrt{\frac{n}{Cx}}$, and $\frac{\sqrt{x}}{\sqrt{n}}\xi_k\le \frac{\sqrt{x}}{\sqrt{n}}\frac{2^{(l+1)/2}}{\sqrt2-1}\le\frac{\sqrt{x}}{\sqrt{n}}\sqrt{\frac{n}{Cx}}\frac{\sqrt2}{\sqrt2-1}=\frac{\sqrt2}{(\sqrt2-1)\sqrt{C}}$. Hence $\exp\left\{\frac{\sqrt{x}}{\sqrt{n}}\xi_k\right\}\le 1+\bar{C}\frac{\sqrt{x}}{\sqrt{n}}\xi_k$ with an appropriate number $\bar{C}=\bar{C}(C)>0$, and $E\exp\left\{\frac{\sqrt{x}}{\sqrt{n}}\xi_k\right\}\le 1+\bar{C}\frac{\sqrt{x}}{\sqrt{n}}E\xi_k\le 1+\bar{K}\frac{\sqrt{x}}{\sqrt{n}}$ with an appropriate number $\bar{K}=\bar{K}(C)$, as we have claimed, since $E\xi_k=\sum\limits_{j=1}^{l}2^{-j/2-1}\le\sqrt2+1$.

This implies that

$$P\left(\frac{\sqrt{72K}}{\sqrt{n}}\sum_{k=1}^{B\sqrt{nx}}\xi_k>\sqrt{x}\right)=P\left(\exp\left\{\frac{\sqrt{x}}{\sqrt{n}}\sum_{k=1}^{B\sqrt{nx}}\xi_k\right\}>\exp\left\{\frac{x}{\sqrt{72K}}\right\}\right)$$
$$\le\left(1+\bar{K}\frac{\sqrt{x}}{\sqrt{n}}\right)^{B\sqrt{nx}}\exp\left\{-\frac{x}{\sqrt{72K}}\right\}\le e^{B\bar{K}x-x/\sqrt{72K}}.$$

Let us choose the number $B=\dfrac{1}{12\bar{K}\sqrt{K}}$ in the last inequality. In such a way we have proved that the probability considered in this relation is less than $e^{-\text{const.}\,x}$. On the other hand, by the first statement of problem 10 the inequality $P(|\kappa_n-E\kappa_n|>B\sqrt{nx})<e^{-\text{const.}\,x}$ also holds. These estimates together with relation (13) imply formula (14).

12.) By summing up the inequalities in formula (15) for $j=1,\dots,l$ (with the coefficient $B=5$) we get that

$$72K\sum_{j=1}^{l}2^jX_n\left(\frac{1}{2^{j-1}}\right)^2$$
$$\le 360K\sum_{j=1}^{l}2^j\left(\sum_{k=j}^{l}2^{(k-j)/2}\left[X_n\left(\frac{1}{2^{k-1}}\right)-X_n\left(\frac{1}{2^k}\right)\right]^2+2^{(l-j)/2}X_n^2\left(\frac{1}{2^l}\right)\right)$$

$$= 360K \sum_{k=1}^{l} 2^k \left[X_n\left(\frac{1}{2^{k-1}}\right) - X_n\left(\frac{1}{2^k}\right)\right]^2 \sum_{j=1}^{k} 2^{(j-k)/2} + 360K \sum_{j=1}^{l} 2^{(l+j)/2} X_n^2\left(\frac{1}{2^l}\right)$$

$$\le 1500K \left(\sum_{k=1}^{l} 2^k \left[X_n\left(\frac{1}{2^{k-1}}\right) - X_n\left(\frac{1}{2^k}\right)\right]^2 + 2^l X_n^2\left(\frac{1}{2^l}\right)\right).$$

The random variables $2^k\left(X_n\left(\frac{1}{2^{k-1}}\right) - X_n\left(\frac{1}{2^k}\right)\right)$, $1 \le k \le l$, and $2^l X_n\left(\frac{1}{2^l}\right)$ are independent, and their joint distribution agrees with the joint distribution of the random variables $\dfrac{\eta_k - E\eta_k}{\sqrt{n}}$, $1 \le k \le l+1$, where η_k, $1 \le k \le l+1$, are those Poissonian random variables which appear in formula (16). Hence the last inequality implies formula (16).

The inequality $P(|\eta_k - E\eta_k| > u) \le 2\exp 2\left\{-\dfrac{u^2}{8n2^{-k}}\right\}$ was already proved in problem 10 in the case $u < n2^{-k}$. Indeed, the estimate proved for a Poissonian random variable κ_n with parameter n holds for all real (not necessary integer) parameter $n > 0$. (In these inequalities we have not tried to give estimates with optimal constants. Further we have formulated them in such a way that the case $k = l+1$ has not to be considered separately.) In particular, with the choice $u = n2^{-k}$ we get the estimate $P(|\eta_k - E\eta_k| \ge n2^{-k}) \le 2\exp\left\{-\dfrac{n}{2^{(k+3)}}\right\}$. To prove the inequality $E\exp\left\{n2^{-(k+4)}\bar\eta_k^2\right\} \le B$ let us introduce the distribution functions $F_k(y) = F_{n,k}(y) = P(|\bar\eta_k| < y)$, $1 \le k \le l+1$, and let us observe that by the results already proved $1 - F_k(y) \le 2e^{-2^k y^2/8n}$. Hence we get by integrating by parts that

$$\begin{aligned}
E\exp\left\{\frac{2^{k-4}}{n}\bar\eta_k^2\right\} &= \int_0^{2^{-k}n} e^{2^k y^2/16n} F_k(\,dy) \\
&= \int_0^{2^{-k}n} (1 - F_k(y))\, de^{2^k y^2/16n} - \left[(1 - F_k(y))e^{2^k y^2/16n}\right]_0^{2^{-k}n} \\
&= \int_0^{2^{-k}n} (1 - F_k(y)) \frac{2^k y}{8n} e^{2^k y^2/16n}\, dy + 1 - (1 - F_k(2^{-k}n))e^{2^{-k}n/16} \\
&\le 2\int_0^{2^{-k}n} \frac{2^k y}{8n} e^{-2^k y^2/16n}\, dy + 1 - e^{-2^{-k}n/16} \\
&= 2\int_0^{2^{-k/2}n^{1/2}/4} 2ye^{-y^2}\, dy + 1 - e^{-2^{-k}n/16} \le B,
\end{aligned}$$

as we have claimed.

In the proof of formula (17) we use formula (16), the truncation of the random variables $\eta_k - E\eta_k$ introduced before, the definition of the random variables $\bar\eta_k$ and the already proved inequalities for the random variables $\eta_k - E\eta_k$. We get with their help that

$$
\begin{aligned}
P\left(72K\sum_{j=1}^{l} 2^j X_n\left(\frac{1}{2^{j-1}}\right)^2 > x\right) &\le P\left(1500K\left(\sum_{k=1}^{l}\frac{2^k}{n}\bar\eta_k^2+\frac{2^l}{n}\bar\eta_{l+1}^2\right)>x\right)\\
&\qquad+\sum_{k=1}^{l+1} P\left(|\eta_k - E\eta_k| > 2^{-k}n\right)\\
\le P\left(\exp\left\{\sum_{k=1}^{l}\frac{2^{(k-4)}}{n}\bar\eta_k^2+\frac{2^{(l-4)}}{n}\bar\eta_{l+1}^2\right\}>\exp\left\{\frac{x}{24000K}\right\}\right)&\\
+2\sum_{k=1}^{l+1} e^{-n2^{-(k+3)}} \le e^{Bl-x/24000K}+\text{const.}\, e^{-n2^{-(l+4)}}.&
\end{aligned}
$$

If $x > C_0 \log n$ with a sufficiently large constant $C_0 > 0$, then $Bl - \dfrac{x}{24000K} \le -\dfrac{x}{30000K}$ and $n2^{-(l+4)} \ge \text{const.}\, x$, since $Bl \le B\log n + \text{const.} \le \dfrac{B}{C_0}x + \text{const.} \le \dfrac{x}{120000}$ under these conditions with a sufficiently large constant C_0, and since $n2^{-l} \ge Cx$, hence $n2^{-(l+4)} \ge \text{const.}\, x$. These inequalities imply formula (17).

Inequality (10) is a simple consequence of formulas (12), (14) (17), and relation (9) holds because of the result of problem 9 from relation (10) and the results of problems 7 and 8.

13.) Both a Wiener process $W(t)$ and a standardized Poisson process $X_n(t)$ are processes with independent increments, and the identities $EW(t) = 0$ and $EX_n(t) = 0$ hold for all numbers $0 \le t \le 1$. Furthermore, the trajectories of a Wiener process are continuous and the trajectories of a standardized Poisson process are cadlag (continuous from the right with left-hand side limits) functions. Hence the *Lemma* formulated in this work can be applied both for the processes $\pm B(t)$ and $\pm X_n(t)$.

Since $W(t)$, $t \ge 0$, is a normally distributed random variable with expectation zero and variance t, hence $Ee^{\pm sW(t)} = e^{ts^2/2}$, and by the last inequality of the *Lemma*

$$
P\left(\sup_{0\le t<L\frac{y}{n}} \pm\sqrt{n}W(t) > y\right) \le \exp\left\{ns^2L\frac{y}{2n} - sy\right\}
$$

with arbitrary number $s > 0$. With the choice $s = \frac{1}{L}$ this inequality yields the analog of inequality (18a) for the Wiener process with parameter $\alpha = \frac{1}{2L}$.

The proof of the analog of inequality (18b) for a standardized Poisson process is similar. The random variable $\sqrt{n}X_n(t)$ is a Poisson distributed random variable with parameter $\sqrt{n}t$ minus its expected value. Hence, as it was shown for instance in the solution of problem 10, $Ee^{\pm s\sqrt{n}X_n(t)} \le e^{s^2 t}$, if $0 \le s \le 1$. Hence an application of the lemma yields with the choice $s = \frac{1}{2L}$ that

$$P\left(\sup_{0\le t<L\frac{y}{n}} \pm\sqrt{n}X_n(t) > y\right) \le e^{s^2Ly-sy} = e^{-y/4L},$$

if $L \ge \frac{1}{2}$, and $s \le 1$ as a consequence. If $L \le \frac{1}{2}$, then exploiting that the probability at the left-hand side of formula (18) is a monotone increasing function of the parameter L we get that the estimate (18b) holds in this case with the same coefficient as in the case $L = \frac{1}{2}$. (Actually the inequality could be improved in this case, but we shall not need such an improvement.)

14.) The proof of formula (18a) can be obtained from the result of problem 13 by means of the representation of a Brownian bridge through a Wiener process in the following way:

$$P\left(\sup_{0\le t<L\frac{y}{n}} \sqrt{n}|B(t)| > y\right) \le P\left(\sup_{0\le t<L\frac{y}{n}} \sqrt{n}|W(t)| > \frac{y}{2}\right)$$
$$+ P\left(\sqrt{n}L\frac{y}{n}|W(1)| \ge \frac{y}{2}\right) \le 2e^{-\alpha y} + P\left(|W(1)| \ge \frac{\sqrt{n}}{2L}\right) \le 2e^{-\alpha y} + 2e^{-n/8L^2}.$$

Formula (18a) follows from this inequality because of the condition $0 < y \le n$.

Formula (18b) can be proved similarly with the help of the Poisson approximation defined in formulas (11a)—(11c) on the basis of the result of problem 13. This yields that

$$P\left(\sup_{0\le t<L\frac{y}{n}} \sqrt{n}|Z_n(t)| > y\right) \le 2e^{-\alpha y} + P\left(\sup_{0\le t<L\frac{y}{n}} \sqrt{n}|Y_n(t)| > \frac{y}{2}\right), \tag{2.4}$$

where the series of random variables $Y_n(t)$ are defined in formula (11c). The second term at the right-hand side of formula (2.4) can be bounded similarly to the method applied in problem 11. Let us consider the conditional distribution of the process $Y_n(t)$ under the condition $\kappa_n = m$, $m = 0, \pm1, \pm2, \dots$. Then we get similarly to the proof of formula (2.3) that

$$P\left(\sup_{0\le t<L\frac{y}{n}} \sqrt{n}|Y_n(t)| > \frac{y}{2}\right) \le P\left(\sup_{0\le t<L\frac{y}{n}} \sqrt{n}|\bar{Y}_{n,B\sqrt{ny}}(t)| > \frac{y}{2}\right) + P\left(|\kappa_n - n| > B\sqrt{ny}\right), \tag{2.5}$$

where the process $\bar Y_{n,m}(t)$ was defined before the formulation of problem 10, before formula (13), and κ_n is the Poisson distributed random variable with parameter n which appears in the definition of the process $Y_n(t)$ given in formula (11), and $B>0$ is an arbitrary positive number. The right-hand side of formula (2.5) can be well estimated by means of the calculation

$$P\left(\sup_{0\le t<L\frac yn}\sqrt n|\bar Y_{n,B\sqrt{ny}}(t)|>\frac y2\right)=P\left(\sqrt n\bar Y_{n,B\sqrt{ny}}\left(\frac{Ly}{n}\right)\ge\frac y2\right)$$
$$=P\left(\sum_{j=1}^{B\sqrt{n}y}\chi_j>\frac y2\right)\le\left(Ee^{\chi_1}\right)^{B\sqrt{ny}}e^{-y/2},$$

where $\chi_1,\chi_2,\dots,$ are independent, identically distributed random variables with distribution $P(\chi_1=1)=1-P(\chi_1=0)=L\frac yn$. Hence we get, because of the condition $y\le n$ that $Ee^{\chi_1}=1+\frac{Ly}{n}(e-1)\le\exp\left\{(e-1)\frac{Ly}{n}\right\}\le\exp\left\{L(e-1)\sqrt{\frac yn}\right\}$.

Let us apply the previous estimates with the choice of parameter $B=\frac1{6L}$ together with the estimate on the distribution function κ_n-n given in problem 10. They yield the inequalities

$$P\left(\sup_{0\le t<L\frac yn}\sqrt n|\bar Y_{n,B\sqrt{ny}}(t)|>\frac y2\right)\le\left(Ee^{\chi_1}\right)^{\sqrt{ny}/6L}e^{-y/2}\le e^{(e-1)y/6-y/2}\le e^{-y/6},$$

and $P\left(|\kappa_n-n|>\frac1{6L}\sqrt{yn}\right)\le e^{-\text{const.}\,y}$. These estimates yield a bound on the expression in formula (2.5) with the choice $B=\frac1{6L}$ which implies that the expression in formula (2.4) is less than $2e^{-\alpha y}$ with an appropriate constant $\alpha>0$. In such a way we have solved problem 14.

15.) Let us choose such numbers $C_0>0$, $C>0$ and $D>0$ and threshold index n_0 for which relation (9) holds if the real number $x>0$ and integer $l>0$ satisfy the conditions $C_0\log n\le x\le C^{-1}n$ and $2^{-l}\ge Cxn^{-1}$, and $n\ge n_0$. First we show the slightly weaker result by which the estimate of the *Approximation Theorem* holds for all numbers x which satisfy the relation $C_0\log n\le x\le C^{-1}n$ with these numbers C and C_0 if $n\ge n_0$.

Let us choose that positive integer $l=l(x)$ for which $2Cxn^{-1}>2^{-l}\ge Cxn^{-1}$ with the constant $C>0$ considered above. We show with the help of relations (18a) and (18b) proved in problem 14 that

$$P\left(\sup_{1\le k\le 2^l}\sup_{(k-1)2^{-l}\le t<k2^{-l}}\sqrt n\left|B(t)-B\left(\frac{(k-1)}{2^l}\right)\right|>\frac x4\right)\le e^{-\alpha x},$$
$$P\left(\sup_{1\le k\le 2^l}\sup_{(k-1)2^{-l}\le t<k2^{-l}}\sqrt n\left|Z_n(t)-Z_n\left(\frac{(k-1)}{2^l}\right)\right|>\frac x4\right)\le e^{-\alpha x}\tag{2.6}$$

with an appropriate constant $\alpha > 0$ and the previously defined integer $l = l(x)$ if $x \ge C_0 \log n$ with an appropriate constant $C_0 > 0$. To prove relation (2.6) let us observe that formulas (18a) and (18b) remain valid if at their left-hand side the domain $0 \le t \le L\frac{y}{n}$ where supremum is taken is replaced by another domain $u \le t \le u + L\frac{y}{n}$ such that $0 \le u \le 1 - \frac{y}{n}$. Indeed, the probability of the event obtained after such a replacement agrees with the original probability. We show with the help of this modified version of formulas (18a) and (18b) with the choice of parameters $y = \frac{x}{4}$, $L = \frac{8}{C}$ and $u = (k-1)2^{-l}$, $1 \le k \le 2^l$ that the probabilities at the left-hand side of the expression (2.6) are less than $2^{l+1}e^{-\alpha x}$. To show this it is enough to check that the inner supremum in those expressions were taken on intervals of length $2^{-l} \le 2\frac{x}{Cn} = L\frac{x}{4n}$ and the outside supremum is taken over 2^l terms. Let us finally observe that $2^{l+1} \le \frac{2n}{Cx} \le n \le e^{\alpha x/2}$, if $x \ge C_0 \log n$, i.e. if $n \le e^{x/C_0}$ with a sufficiently large number C_0. This argument implies formula (2.6) (with parameter $\alpha/2 > 0$ instead of parameter $\alpha > 0$.)

The above mentioned weakened form of the *Approximation Theorem* is a simple consequence of formulas (9) and (2.6). Indeed, given a number $0 \le t \le 1$, let us consider the integer $k = k(t)$, $1 \le k \le 2^l$ for which $(k-1)2^{-l} \le t < k2^{-l}$. Then we get by formulas (9) and (2.6) that in the case $x \le C_0 \log n$ the inequality

$$\begin{aligned}\sqrt{n}\,|Z_n(t) - B(t)| &\le \sqrt{n}\left|Z_n\left((k-1)2^{-l}\right) - B\left((k-1)2^{l}\right)\right| \\ &\quad + \sqrt{n}\left|B(t) - B\left((k-1)2^{-l}\right)\right| + \sqrt{n}\left|Z_n(t) - Z_n\left((k-1)2^{-l}\right)\right| \le x\end{aligned}$$

holds simultaneously for all numbers $0 \le t \le 1$ except on a set of probability $e^{-Dx} + 2e^{-\alpha x}$. Hence the approximation theorem holds for $C_0 \log n < x \le C^{-1}n$. The estimate of the *Approximation Theorem* is a trivial statement in the case $x \le C_0 \log n$ if the constant C_1 in it is chosen sufficiently large.

In the case $x \ge C^{-1}n$ we can show with the help of formulas (18a) and (18b) that

$$\begin{aligned}&P\left(\sqrt{n}\sup_{0\le t\le 1}|Z_n(t) - B(t)| > x\right) \\ &\quad \le P\left(\sqrt{n}\sup_{0\le t\le 1}|B(t))| > \frac{x}{2}\right) + P\left(\sqrt{n}\sup_{0\le t\le 1}|Z_n(t))| > \frac{x}{2}\right) \\ &\quad \le e^{-Cx^2/n} \le e^{-\bar{C}x}\end{aligned} \tag{2.7}$$

with appropriate constants $C > 0$ and $\bar{C} > 0$. Let us remark that the condition $0 \le \frac{x}{2} \le n$ appears in formulas (18a) and (18b), hence we need some special considerations in the proof. The first term in the second line of formula (2.7) can be estimated similarly to the expression in formula (18a) for all numbers $x > C^{-1}n$ which expression yields an estimate on the supremum of a Brownian bridge. (In the proof of this estimate we applied the representation of a Brownian bridge by a Wiener process, and the condition $x \le n$ was not needed there. Now we wrote a sharp upper bound for the probability we wanted to estimate, only the constant $C > 0$ appearing there is not calculated explicitly.) The Poisson approximation applied in the proof of formula (18b) does not yield a good estimate in the case $x \gg n$, but in the case $\frac{x}{2} > n$ we do not need this approximation. In this case the trivial identity

$$P\left(\sqrt{n}\sup_{0\le t\le 1}|Z_n(t))| > n\right) = 0$$

is applicable. Hence formula (2.7) holds, and the *Approximation Theorem* holds for all $x \ge C^{-1}n$.

Although the special investigation of the case $n \le n_0$ has no great importance, we remark that the *Approximation Theory* holds in this case, too. Too see this it is enough to observe that since the parameter n is bounded, hence the second, and as a consequence the first line of inequality (2.7) is less than $C_1e^{-C_2x}$ for all numbers $x \ge 0$ with some appropriate constants $C_1 > 0$ and $C_2 > 0$.

Appendix

Proof of the Lemma:

Let us define the following random variable (stopping rule) τ which tells us the smallest index k for which an element of the sequence of partial sums $S_k = \sum\limits_{j=1}^{k} \xi_k$, $k = 1, \dots, n$, is larger than a number $x > 0$:

$$\tau = \tau(x, n) = \begin{cases} \min\{k\colon S_k > x\} & \text{if } \sup\limits_{1 \le k \le n} S_k > x \\ n & \text{if } \sup\limits_{1 \le k \le n} S_k \le x \end{cases}.$$

As $P\left(\sup\limits_{1 \le k \le n} S_k > x\right) = P\left(S_\tau > x\right)$, it is natural to prove the first inequality of the *Lemma* by giving a good estimate on the exponential moments Ee^{sS_τ} of the random variable S_τ.

It follows from standard results of the martingale theory that for arbitrary real number $s \ge 0$ $Ee^{sS_\tau} \le Ee^{sS_n}$. To prove this it is useful to show that for all numbers $s \ge 0$ the sequence $(e^{sS_k}, \mathcal{F}_k)$, $k = 1, \dots, n$, where $\mathcal{F}_k = \sigma(\xi_1, \dots, \xi_k)$ is the σ-algebra generated by the random variables $\xi_1, \dots, \xi_k$, is a supermartingale i.e. $E\left(e^{sS_{k+1}} | \mathcal{F}_k\right) \ge e^{sS_k}$ with probability one. It is not difficult to see this inequality by using the properties of conditional expectations, the identity $e^{sS_{k+1}} = e^{sS_k} e^{s\xi_{k+1}}$ and the independence of the random variable ξ_{k+1} of the σ-algebra $\mathcal{F}_k$. This implies that $E\left(e^{sS_{k+1}} | \mathcal{F}_k\right) = e^{sS_k} Ee^{s\xi_{k+1}} \ge e^{sS_k}$, since $Ee^{s\xi_{k+1}} \ge e^{Es\xi_{k+1}} \ge 1$ by the Jensen inequality and the condition $E\xi_{k+1} \ge 0$.

A simple but fundamental result of the martingale theory states that if the series $(e^{sS_k}, \mathcal{F}_k)$, $k = 1, \dots, n$, is a supermartingale and the stopping rule τ satisfies the inequality $\tau \le n$ with probability one, then $Ee^{sS_\tau} \le Ee^{sS_n}$. It is not difficult to prove the above inequality, but since the argument of the proof is essentially different from the method of this work we omit it. It may be worth mentioning that this inequality has the following heuristic content: In an advantageous game the further we play the greater our expected gain will be.

It follows from the (only partly proved) estimate on the exponential moment that

$$\begin{aligned} P\left(\sup_{1 \le k \le n} S_k > x\right) &= P\left(S_\tau > x\right) = P\left(e^{sS_\tau} > e^{sx}\right) \le Ee^{sS_\tau} e^{-sx} \\ &\le Ee^{sS_n} e^{-sx} = \exp\left\{-sx + \sum_{k=1}^{n} B_k(s)\right\}, \end{aligned}$$

and this is the first statement of the Lemma.

To prove the second statement of the lemma let us introduce for all numbers $n = 1, 2, \dots$ the numbers $t_{k,n} = a + (b - a)k2^{-n}$, $0 \le k \le 2^n$ and the random variables

$\xi_{k,n} = X(t_{k,n}) - X(t_{k-1,n})$, $1 \le k \le 2^n$, Then for a fixed number n the random variables $\xi_{k,n}$, $1 \le k \le 2^n$, are independent, since the process $X(t)$ has independent increments and $X(t_{k,n}) - X(a) = \sum\limits_{j=1}^{k} \xi_{j,n}$, $1 \le k \le n$. Furthermore, $E\xi_{k,n} \ge 0$ for $1 \le k \le n$, and since the trajectories of the stochastic process $X(t)$ are continuous or cadlag (continuous from the right with left-hand side limit) functions, hence for all numbers $x > 0$

$$\left\{\omega\colon \sup_{a\le t\le b} (X(t,\omega) - X(a,\omega)) > x\right\} = \bigcup_{n=1}^{\infty} \left\{\omega\colon \sup_{1\le k\le 2^n} (X(t_{k,n},\omega) - X(a,\omega)) > x\right\}.$$

Beside this the sets in the union at the right-hand side of the last relation constitute a series of sets monotone increasing by the parameter n. From this fact and the already proven part of the *Lemma*

$$P\left(\sup_{a\le t\le b} (X(t) - X(a)) > x\right) = \lim_{n\to\infty} P\left(\sup_{1\le k\le 2^n} (X(t_{k,n}) - X(a)) > x\right)$$
$$\le \lim_{n\to\infty} e^{-sx} \prod_{k=1}^{2^n} Ee^{s\xi_{k,n}} = e^{-sx} Ee^{s(X(b)-X(a))},$$

and this is the second statement of the *Lemma.*