Approximation of partial sums of independent random variables

Let $X_1, X_2, \dots$ be a sequence of independent and identically distributed random variables, and let us consider the partial sums $S_0 = 0$, $S_n = \sum\limits_{k=1}^n X_k$, $n = 1, 2, \dots$, defined by them. Let us also define the following $S(t)$, $0 \le t < \infty$, (random) broken line:

$$\begin{aligned} &S(n) = S_n, \quad n = 0, 1, 2, \dots, \\ &S(t) = S(n) + (t - n)(S(n+1) - S(n)) \ \text{ if } \ n \le t < n+1, \quad n = 0, 1, 2, \dots. \end{aligned} \tag{1}$$

The stochastic process $S(t)$, $t \ge 0$, behaves similarly to a Wiener process $W(t)$, $t \ge 0$, (i.e. to a Gaussian stochastic process $W(t)$ with continuous trajectories with expectation $EW(t) = 0$ for all $t \ge 0$ and covariance function $EW(s)W)t) = \min(s, t)$ for all pairs of numbers $0 \le s, t < \infty$). In this work we are interested in the question how well the process $S(t)$ can be approximated by means of an appropriate Wiener process. For the sake of convenience we shall study the following equivalent problem. Given a Wiener process $W(t)$, $0 \le t < \infty$, at the start we want to construct a broken line type random process $\bar S(t)$, $0 \le t \le \infty$, with the same distribution as the random broken line $S(t)$, $0 \le t < \infty$, defined in (1) which is as close to the Wiener process $W(t)$ as possible. More explicitly, we want to get such a construction for which the probability $P\left(\sup\limits_{0 \le t \le T} |\bar S(t) - W(t)| > A(T)\right)$ is almost zero for all sufficiently large parameters T, and we would like to have this relation with a function $A(T)$, $T \ge 0$, as small as possible. This question is a natural counterpart of the problem studied in the series of problems *The approximation of the normalized empirical distribution function by a Brownian bridge.* A similar result can be proved also in this case. Namely, the following Theorem holds.

Approximation Theorem. *Let F be a distribution function such that*

$$\int x F(\,dx) = 0, \quad \int x^2 F(\,dx) = 1, \quad \int e^{sx} F(\,dx) < \infty, \quad \text{if } |s| < s_0 \tag{2}$$

with some appropriate number $s_0 > 0$, (i.e. a random variable X with distribution function F satisfies the relations $EX = 0$, $EX^2 = 1$ and $Ee^{sX} < \infty$ if the absolute value of the number s is small). Let a Wiener process $W(t) = W(t, \omega)$, $t \ge 0$, be given on some probability space $(\Omega, \mathcal{A}, P)$. Then a sequence of independent, identically distributed random variables $X_1, X_2, \dots$ with distribution function F can be constructed on this probability space $(\Omega, \mathcal{A}, P)$ in such a way that the random broken line $S(t) = S(t, \omega)$, $t \ge 0$, defined by means of the partial sums $S_0 = 0$, $S_n = \sum\limits_{k=1}^n X_k$, $n = 1, 2, \dots$, in formula (1) satisfies the inequality

$$P\left(\sup_{0 \le t \le T} |S(t, \omega) - W(t, \omega)| > C_1 \log T + x\right) < C_2 e^{-\lambda x} \tag{3}$$

for all numbers $x \ge 0$ with some appropriate constants $C_1 > 0$, $C_2 > 0$ and $\lambda > 0$ depending only on the distribution function F.

It is not difficult to prove the following statement of Problem 1.

1.) If the random broken line $S(t) = S(t,\omega)$ and Wiener process $W(t) = W(t,\omega)$, $t \ge 0$ satisfy relation (3), then there exists some constant $K > 0$ such that

$$\limsup_{T\to\infty} \frac{\sup\limits_{0\le t\le T} |S(t,\omega) - T(t,\omega)|}{\log T} < K \quad \text{with probability 1.} \tag{4}$$

The approximation theorem or its consequence formulated in Problem 1 states in a slightly informal interpretation that in the case of an appropriate construction the relation $|S(t) - W(t)| = O(\log n)$ holds. On the other hand, the order of magnitude of the random variables $S(t)$ and $W(t)$ is const. $\sqrt{t}$. This means that the estimation of the Approximation theorem yields an approximation of the same order as the result of the series of problems *The approximation of the normalized empirical distribution function by a Brownian bridge* for the approximation of the normalized empirical distribution function by a Brownian bridge. I shall formulate two Statements whose content is that the estimate of the Approximation theorem is sharp. More explicitly, there is no such construction for the approximation of partial sums by a Wiener process which would yield a version of formula (4) with some function $g(T)$ such that $g(T) = o(\log T)$ as $T \to \infty$. Beside this, condition (2) of the Approximation Theorem cannot be dropped. In more detail, I formulate the following results:

Statement 1. *Let $X_1, X_2, \dots,$ be a sequence of independent and identically distributed random variables on a probability space $(\Omega, \mathcal{A}, P)$. Put $S_0 = 0$, $S_n = \sum\limits_{k=1}^{n} X_k$, $n = 1, 2, \dots,$ and define the random broken line $S(t) = S(t,\omega)$ from these random variables by means of formula (1). Let $W(t)$, $t \ge 0$, be a Wiener process on the same probability space $(\Omega, \mathcal{A}, P)$. If the random variables X_k are not standard normal distributed, then there exists some constant $K > 0$ such that*

$$\limsup_{T\to\infty} \frac{\sup\limits_{0\le t\le T} |S(t,\omega) - T(t,\omega)|}{\log T} > K \quad \textit{with probability 1.} \tag{5}$$

Statement 2. *If the random broken line $S(t)$, $t \ge 0$, made from the partial sums of some independent and identically distributed random variables $X_1, X_2, \dots$ by means of formula (1) and a Wiener process $W(t)$, $t \ge 0$ satisfy relation (4), then $EX_1 = 0$, $EX_1^2 = 1$ and $Ee^{sX_1} < \infty$ for $|s| < s_0$ with some number $s_0 > 0$, i.e. the distribution function $F(x)$ of the random variable X_1 satisfies relation (2).*

Statements 1 and 2 follow from the results of Problems 10 and 12 of this note.

Not only the result of the Approximation Theorem is similar to the main result of the series of problems in *The approximation of the normalized empirical distribution function by a Brownian bridge*, but also the constructions yielding these results are based on a similar idea. To explain the relation between them it is useful to formulate the version of the Approximation Theorem about a sequence of finitely many partial sums.

The finite version of the Approximation Theorem. *Let us fix a positive integer n. Let F be a distribution function satisfying the properties given in formula (2), and let $W(t)$, $0 \le t \le 2^n$, be a Wiener process on the interval $0 \le t \le 2^n$. A sequence of independent and identically distributed random variables X_k, $1 \le k \le 2^n$, can be constructed with distribution function F in such a way that the partial sums $S_0 = 0$, $S_k = \sum\limits_{j=1}^{k} X_j$, $1 \le k \le 2^n$, and random broken line function $S_n(t)$, $0 \le t \le 2^n$ defined by the formula*

$$S_n(t) = S_{k-1} + (t-(k-1))(S_k - S_{k-1}), \quad \textit{if } k-1 \le t \le k, \quad 1 \le k \le 2^n, \tag{1'}$$

satisfy the following inequalities:

$$P\left(\sup_{0 \le t \le 2^n} |S_n(t,\omega) - W(t,\omega)| > \bar C_1 n + x\right) < \bar C_2 e^{-\lambda x}, \tag{6}$$

and in the end point $t = 2^n$

$$P\left(|S_{2^n}(\omega) - W(2^n,\omega)| \ge C_1 + x\right) \le \bar C_2 e^{-\lambda x} \tag{6a}$$

for all numbers $x > 0$ with some appropriate constants $\bar C_1 > 0$, $\bar C_2 > 0$ and $\lambda > 0$ depending only on the distribution function F.

Remark. *Relation (6a) can be proved as the consequence of the following statement. A sequence of independent random variables satisfying the finite version of the Approximation Theorem for which*

$$|S_{2^n}(\omega) - W(2^n,\omega)| \le C\frac{W(2^n,\omega)^2}{2^n} + D, \quad \textit{if } |W(2^n,\omega)| \le \varepsilon 2^{n/2}$$

with some appropriate constants $C > 0$, $D > 0$ and $\varepsilon > 0$.

Let us prove the following (simple) statement.

2.) The Approximation Theorem can be deduced from the *Finite version of the Approximation Theorem.*

A construction leading to the proof of the *Finite version of the Approximation Theorem* can be obtained as a natural adaptation of the construction described in *The approximation of the normalized empirical distribution function by a Brownian bridge*, at least if we impose some additional conditions about the distribution function F in this result.

Given a Wiener process $W(t)$, $0 \le t \le 2^n$, first we construct the random variable S_{2^n} as the quantile transform of $W(2^n)$, i.e. let $S_n(2^n) = S_{2^n} = F_{2^n}^{-1}\left(\Phi\left(\frac{W(2^n)}{2^{n/2}}\right)\right) = \left\{u\colon F_{2^n}(u) < \Phi\left(\frac{W(2^n)}{2^{n/2}}\right)\right\}$, where $F_{2^n}(x) = \bar F_{2^n}(2^{n/2}x)$, and $\bar F_{2^n}(x)$ is the distribution function of the random variable S_{2^n}, i.e. it equals the 2^n-times convolution of the function F with itself, and $\Phi(x)$ is the standard normal distribution function, hence it is the distribution function of the random variable $\frac{W(2^n)}{2^{n/2}}$. Since the distribution function of the sum of N independent, identically distributed random variables with expectation zero and variance 1 can be well approximated by the normal distribution function with expectation zero and variance N, and also a good large deviation type result is known about this approximation (this result is also proved in Problem 22 of the series of problems *The theory of large deviations I.* which exists only in Hungarian for the time being), it can be proved with the help of some calculation that the above construction satisfies relation (6a). In the approximation of the standardized empirical distribution function by a Brownian bridge $B(t)$ no step corresponding to this argument appears, since $Z_n(1) = B(1) = 0$, in the end-point $t = 1$. Hence in the point $t = 1$ the random process $Z_n(t)$ need not be fitted to the process $B(t)$.

After the definition of the value of the stochastic process $S_n(t)$ in the end-points $t = 2^n$ and $t = 0$ (we have $S_n(0) = 0$), we can define its values in the points $t = (2k-1)2^{n-l}$, $1 \le k \le 2^{l-1}$, by means of induction with respect to the parameter l, $l = 1, \dots, n$, as an appropriate transform of the Wiener process $W(t)$, $0 \le t \le 2^n$. This definition is a natural adaptation of the "halving" construction of the normalized empirical distribution function by means of the Brownian bridge described in *The approximation of the normalized empirical distribution function by a Brownian bridge.* The main difference between these constructions is that now the Brownian bridge $B(t)$ is replaced by the Wiener process $2^{-n/2}W(2^n t)$ and the normalized empirical distribution function $Z_n(t)$ by the random broken line $2^{-n/2}S_n(2^n t)$, $0 \le t \le 1$. More explicitly, our definition is based on the following observation.

If $S_1, \dots, S_{2^n}$ are partial sums of independent, identically distributed random variables with distribution F, and $W(t)$, $0 \le t \le 2^n$, is a Wiener process, then we define the analogs of the random variables $U_{k,l}$, $V_{k,l}$, $\bar U_{k,l}$ and $\bar V_{k,l}$ and σ-algebras $\mathcal{F}_l$ and $\mathcal{G}_l$ introduced in formulas (1)–(4) of *The approximation of the normalized empirical distribution function by a Brownian bridge* by means of the following formulas:

$$U_{k,l} = U_{k,l,n} = 2^{(l-n+1)/2}\left[W\left(k2^{n-l}\right) - W\left((k-1)2^{n-l}\right)\right], \quad 1 \le k \le 2^l,\ \ 0 \le l \le n,$$
$$V_{k,l} = V_{k,l,n} = 2^{(l-n+1)/2}\left[S_{k2^{n-l}} - S_{(k-1)2^{n-l}}\right], \quad 1 \le k \le 2^l,\ \ 0 \le l \le n,$$

$$\mathcal{F}_l = \mathcal{B}\left\{U_{k,l},\ 1 \le k \le 2^l\right\}, \quad 0 \le l \le n,$$
$$\mathcal{G}_l = \mathcal{B}\left\{V_{k,l},\ 1 \le k \le 2^l\right\}, \quad 0 \le l \le n,$$

$$\mathbf{U}_l = \{U_{k,l},\ k = 1, \dots, 2^l\}, \quad 0 \le l \le n$$
$$\mathbf{V}_l = \{V_{k,l},\ k = 1, \dots, 2^l\}, \quad 0 \le l \le n$$

and

$$\bar{\mathbf{U}}_{l+1}=\{\bar U_{1,l+1},\dots,\bar U_{2^{l+1},l+1}\},\quad \bar{\mathbf{V}}_{l+1}=\{\bar V_{1,l+1},\dots,\bar V_{2^{l+1},l+1}\},$$
$$\bar U_{k,l+1}=U_{k,l+1}-E(U_{k,l+1}|\mathcal{F}_l),\quad \bar V_{k,l+1}=V_{k,l+1}-E(V_{k,l+1}|\mathcal{G}_l)$$
$$1\le k\le 2^{l+1},\quad 0\le l\le n-1.$$

These random variables $U_{k,l}$, $V_{k,l}$, $\bar U_{k,l}$ and $\bar V_{k,l}$ satisfy the natural analogs of the properties listed in Problems 3 and 4 of *The approximation of the normalized empirical distribution function by a Brownian bridge.* In particular, the following identities hold:

$$\begin{aligned}
\bar V_{2k-1,l+1}&=2^{(l-n)/2}\left(S_{(2k-1)2^{(n-l-1)}}-S_{(k-1)2^{(n-l-1)}}-\frac12\left(S_{k2^{(n-l)}}-S_{(k-1)2^{(n-l)}}\right)\right)\\
&=\frac{2^{(l-n)/2}}{2}\left(\left(S_{(2k-1)2^{(n-l-1)}}-S_{(k-1)2^{(n-l)}}\right)-\left(S_{k2^{(n-l)}}-S_{(2k-1)2^{(n-l-1)}}\right)\right)\\
\bar V_{2k,l+1}&=2^{(l-n)/2}\left(S_{k2^{(n-l)}}-S_{(2k-1)2^{(n-l-1)}}-\frac12\left(S_{k2^{(n-l)}}-S_{(k-1)2^{(n-l)}}\right)\right)\\
&=-\bar V_{2k-1,l+1}
\end{aligned}\tag{7a}$$

and

$$\begin{aligned}
\bar U_{2k-1,l+1}&=2^{(l-n)/2}\Bigg(W((2k-1)2^{(n-l-1)})-W((k-1)2^{(n-l-1)})\\
&\qquad-\frac12\left(W(k2^{(n-l)})-W((k-1)2^{(n-l)})\right)\Bigg)\\
&=\frac{2^{(l-n)/2}}{2}\Bigg(\left(W((2k-1)2^{(n-l-1)})-W((k-1)2^{(n-l)})\right)\\
&\qquad-\left(W(k2^{(n-l)})-W((2k-1)2^{(n-l-1)})\right)\Bigg)\\
\bar U_{2k,l+1}&=2^{(l-n)/2}\Bigg(W(k2^{(n-l)})-W((2k-1)2^{(n-l-1)})\\
&\qquad-\frac12\left(W(k2^{(n-l)})-W(k-1)2^{(n-l)})\right)\Bigg)=-\bar U_{2k-1,l+1}
\end{aligned}\tag{7b}$$

for all numbers $0\le l\le n-1$ and $1\le k\le 2^l$. These relations also imply that $U_{2k-1,l+1}=-U_{2k,l+1}$ for all numbers $1\le k\le 2^l$, and the random variables $U_{2k-1,l+1}$, $0\le l\le n-1$ and $1\le k\le 2^l$ are independent with standard normal distribution.

We have to prove some properties of the random variables $V_{k,l}$ and $\bar V_{k,l}$ introduced in the present investigation which cannot be considered as the natural analogs of the results in *The approximation of the normalized empirical distribution function by a Brownian bridge.* Namely, we have to give a good asymptotic formula for the conditional distribution function of the random variables $V_{k,l+1}$ under the condition of the σ-algebra $\mathcal{G}_l$.

The random variables $\bar V_{2k-1,l+1}$, $1\le k\le l$, are conditionally independent under the σ-algebra $\mathcal{G}_l$ also in the present case. Beside this, the conditional distribution of the random variable $\bar V_{2k-1,l+1}=-\bar V_{2k,l+1}$ under the condition of the σ-algebra $\mathcal{G}_l$ can be expressed explicitly as a function of the random variable $V_{k,l}$. We can write that

$$\begin{aligned} P\left(\bar V_{2k-1,l+1}<x\,\middle|\,\mathcal{G}_l\right)&=P\left(V_{2k-1,l+1}<x\,|\,V_{1,l},\dots,V_{2^l,l}\right)\\ &=P\left(V_{2k-1,l+1}<x\,|\,V_{k,l}\right)=F_{2^{n-l-1}}(x|V_{k,l}),\end{aligned} \tag{8a}$$

where the functions $F_N(x|y)$, $N=1,2,\dots$, are defined by the formula

$$F_N(x|y)=P\left(\sqrt{\frac{2}{N}}\left(S_N-\frac12 S_{2N}\right)<x\,\middle|\,\frac{S_{2N}}{\sqrt{2N}}=y\right),\ \text{where } S_k=\sum_{j=1}^k X_j,\ k=1,2,\dots$$

and $X_1,X_2,\dots,$ are independent random variables with distribution function F (8b)

for all numbers $N=1,2,\dots$.

3.) Let us prove formulas (7a), (7b), (8a) and (8b).

The random variables $X_k=\frac{1}{\sqrt2}V_{k,n}$ can be defined by means of the above formulas similarly to the method of *The approximation of the normalized empirical distribution function by a Brownian bridge* by means of the inductive (with respect to the parameter l) construction by defining first the random variables $\bar V_{2k-1,l}$ and then $V_{2k-1,l}$, $1\le k\le 2^{l-1}$. The only essential difference is that we define the random variable $\bar V_{2k-1,l+1}$ by means of the relation

$$\bar V_{2k-1,l+1}=F^{-1}_{2^{n-l-1}}(\Phi(\bar U_{2k-1,l})|V_{k,l}), \tag{9}$$

where the function $F_N(x|y)$ was defined in formula (8b), and

$$F^{-1}_{2^{n-l-1}}(x|y)=\sup\{u\colon F_{2^{n-l-1}}(u|y)<x\}.$$

(This corresponds to formula (6a) in *The approximation of the normalized empirical distribution function by a Brownian bridge.*

The statement that the above construction satisfies the *Finite version of the Approximation Theorem* can be proved similarly to the corresponding result in *The approximation of the normalized empirical distribution function by a Brownian bridge.* Moreover, since in the present case the partial sums of independent random variables have to be estimated, some steps of the proof become simpler. Here we do not need the Poisson approximation applied in the above mentioned series of problems which was needed to overcome some difficulties arising from the not complete independence of the random variables we had to work with. Let me also remark that in the study of the approximation of the normalized empirical distribution functions we have applied such properties of the binomial and exponential distributions, whose analogs also hold for all random variables with distribution function F satisfying relation (2).

If we consider the approximation result of a random broken line process defined with the help of normalized partial sums of independent random variables in formula (1) by means of a Wiener process in the way described above and want to show that it yields a good approximation one serious problem appears in the proof. We have to show that the random variable $\bar V_{2k-1,l+1}$ defined in formula (9) is sufficiently close to the random variable $\bar U_{2k-1,l+1}$. To do this we need a good estimate on the closeness of conditional distribution function $F_N(x|y)$ defined in formula (8b) to the standard normal distribution function. I shall formulate a sharp estimate for the difference of these two distribution functions under the name *Property A* which will be sufficient for our purposes. But the proof of *Property A* which does not follow directly from standard well-known results is not simple. In the next part I shall concentrate on the proof of this result and the difficulties related to it.

The definition of Property A. *Let a distribution function F be given, and let us consider a sequence X_k, $k=1,2,\dots$ of independent F distributed random variables. Let us define the partial sums $S_n=\sum\limits_{k=1}^n X_k$, $n=1,2,\dots$, of this random variables. We shall say the distribution function F satisfies* Property A *if there exists some number $\varepsilon>0$ and threshold index n_0 such that the relation*

$$\begin{aligned}
1-F_n(x|y)&=P\left(\sqrt{\frac2n}\left(S_n-\frac12S_{2n}\right)>x\,\middle|\,\frac{S_{2n}}{\sqrt{2n}}=y\right)\\
&=(1-\Phi(x))\exp\left\{O\left(\frac{x^3+x^2|y|+|y|+1}{\sqrt n}\right)\right\}\\
&\qquad \textit{if } 0\le x\le\varepsilon\sqrt n,\;\; 0\le|y|\le\varepsilon\sqrt n\\
F_n(-x|y)&=P\left(\sqrt{\frac2n}\left(S_n-\frac12S_{2n}\right)<-x\,\middle|\,\frac{S_{2n}}{\sqrt{2n}}=y\right)\\
&=(1-\Phi(x))\exp\left\{O\left(\frac{x^3+x^2|y|+|y|+1}{\sqrt n}\right)\right\}\\
&\qquad \textit{if } 0\le x\le\varepsilon\sqrt n,\;\; 0\le|y|\le\varepsilon\sqrt n,
\end{aligned}\tag{10}$$

holds, where $\Phi(x)$ is the standard normal distribution function, and the error term $O(\cdot)$ is uniform in the variables x, y and n.

Let me remark that the error term $O(\cdot)$ in formula (10) contains such a polynomial of order 3 which is in his variable $|y|$ only of order 1. An estimate analogous to *Property A* also appeared in *The approximation of the normalized empirical distribution function by a Brownian bridge.* In Problem 2 of that work the approximation of normalized partial sums of independent binomial random variables was considered by standard normal random variables in the case when partial sums of the binomial random variables were divided by a number which might slightly differ from the square root of the variance of this sum. This problem corresponds to the approximation of the random variable defined in formula (9) by standard normal random variable. This problem could be solved with the help of an estimate about the distribution of partial sums of independent random variables with binomial distribution which is a natural analog of Property A.

We shall show that if the distribution function F has moment generating function and a sufficiently smooth density function, then the distribution function F satisfies Property A. We shall prove with the help of the solution of some problems the following Proposition.

Proposition. *Let us assume that the distribution function F satisfies property (2), and its moment generating function $R(s)=\int e^{sx}F(\,dx)$ together with its analytic continuation $R(z)=R(s+it)=\int e^{su+itu}F(\,du)$ satisfies the relation*

$$\int_{-\infty}^{\infty}|R(s+it)|^k\,dt<\infty \tag{11}$$

with some appropriate positive integer $k>0$, if $|s|<s_0$ with some real number $s_0>0$. Then the distribution function F satisfies Property A.

The conditions of the Proposition are satisfied if the distribution function $F(x)$ satisfies Condition (2), and it has a sufficiently smooth density function $f(x)$. In this case the function $R(s+it)$, as a function of the variable t with some fixed number s, is the Fourier transform of the function $e^{sx}f(x)$ which tends to zero as $t\to\infty$ sufficiently fast. Hence Condition (11) is satisfied in this case even with $k=1$.

If the distribution function $F(x)$ satisfies the conditions of the Proposition, then formula (11) guarantees that the distribution function $F(x)$ or of its k-fold convolution with itself has an $f(x)$ or $f_k(x)$) density function whose n-fold convolution with itself is close to the standard normal density function, and there is a good estimate is known for the difference of these density functions. By means of this estimate a good asymptotic formula can be given for the (existing) density function $f_n(x|y)=\frac{\partial}{\partial x}F_n(x|y)$ of the conditional distribution function $F_n(x|y)$ defined in formula (8b). By integrating this density function we can get the proof of the *Proposition.*

In the discussion of the next problems the details of the above method will be worked out. The question may arise whether there exists a different method to prove *Property A.* This question is interesting in particular, because there are such conditions among the conditions of the *Proposition* which do not appear among the conditions of the *Approximation Theorem.* Hence the result of the *Proposition* in its original form may only help to prove a weaker form of the *Approximation Theorem.*

On the other hand I shall also show such an example where a distribution function F satisfies relation (2), but if does not satisfy *Property A.* In this case the construction discussed above is not sufficient for the proof of the *Approximation Theorem.* It will be discussed how to overcome this difficulty.

First I formulate the result about the approximation of the density function of normalized sums of independent, identically distributed random variables by the standard normal density function which we shall apply in the proof of the *Proposition.* This result can be found (with different scaling) in Problem 23 of the series of problems *Theory of*

Large Deviations I. (At present it exists only in Hungarian.) I shall present the proof in the Appendix.

Sharp form of the local central limit theorem. *Let $X_1, X_2, \dots,$ be independent random variables with distribution function F, and put $S_n = \sum\limits_{j=1}^{n} X_j$. Let us assume that the distribution function F satisfies the conditions formulated in relations (2) and (11). Then there exists a number $\varepsilon > 0$ in such a way that the distribution function $F_n(x) = P\left(\frac{S_n}{\sqrt{n}} < x\right)$ has a density function $f_n(x) = \frac{dF_n(x)}{dx}$, and it satisfies the relation*

$$\begin{aligned} f_n(x) &= \exp\left\{\frac{x^3}{\sqrt{n}}\lambda\left(\frac{x}{\sqrt{n}}\right)\right\} \frac{e^{-x^2/2}}{\sqrt{2\pi\left(1+\frac{x}{\sqrt{n}}\mu\left(\frac{x}{\sqrt{n}}\right)\right)}}\left(1+O\left(\frac{1}{\sqrt{n}}\right)\right) \\ &= \varphi(x)\exp\left\{\frac{x^3}{\sqrt{n}}\lambda\left(\frac{x}{\sqrt{n}}\right)\right\}\exp\left\{O\left(\frac{1+|x|}{\sqrt{n}}\right)\right\}, \qquad \textit{if } |x| \le \varepsilon\sqrt{n}, \textit{ and } n \ge k \end{aligned} \tag{12a}$$

(with the number k in formula (11)) where $\mu(x)$ and $\lambda(x)$ are analytic function in a small neighbourhood of the origin, $\varphi(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ is the standard normal density function, and the error term $O(\cdot)$ is uniform in both variables x and n. Also the inequality

$$f_n(x+z) \le \text{const.}\, f_n(x)e^{-sz\sqrt{n}}, \qquad \textit{if } |x| \le \varepsilon\sqrt{n} \textit{ and } n \ge k \tag{12b}$$

holds for arbitrary real number z, where the number s is the solution of the equation $\frac{d}{ds}[\log R(s)] = \frac{x}{\sqrt{n}}$ with $R(s) = \int e^{sx}F(\,dx)$. The const. *in formula (12b) depends only on the distribution function F. Beside this, the above introduced number s satisfies the relation $s \ge 0$ if $x \ge 0$, and $s < 0$ if $x < 0$.*

Beside this, there exists a constant $K > 0$ depending only on the distribution F (but not on the parameter n) such that

$$\sup_{-\infty < x < \infty} f_n(x) \le K \quad \textit{if } n \ge k \tag{12c}$$

with the number k in condition (11).

Let us solve with the help of the above results the following problems.

4.) Let a distribution function F together with a sequence $X_1, X_2, \dots,$ of independent F distributed random variables be given such that for all sufficiently large indices n the distribution function $F_n(x) = P\left(\frac{S_n}{\sqrt{N}} < x\right)$ defined with the help of the partial sums $S_n = \sum\limits_{j=1}^{n} X_j$ has a density function $f_n(x)$. Then the conditional distribution function $F_n(x|y)$ defined in formula (8b) (with the notation of parameter N instead of n in that formula) has a conditional density function $f_n(x|y) = \frac{\partial}{\partial x}F_n(x|y)$ for all large indices n, (the index n can be chosen so large that the density function $f_n(x)$ exists), and it satisfies the identity

$$f_n(x|y) = \frac{f_n\left(\frac{y+x}{\sqrt{2}}\right) f_n\left(\frac{y-x}{\sqrt{2}}\right)}{f_{2n}(y)}. \tag{13a}$$

b) If the distribution function F satisfies relation (12a), then the conditional density function $f_n(y|x)$ satisfies the following version of *Property A*:

$$f_n(x|y) = \frac{e^{-x^2/2}}{\sqrt{2\pi}} \exp\left\{O\left(\frac{|x|^3 + x^2|y| + |y| + 1}{\sqrt{n}}\right)\right\}, \tag{13b}$$

if $0 \le |x| \le \varepsilon\sqrt{n}$, $0 \le |y| \le \varepsilon\sqrt{n}$ with some appropriate number $\varepsilon > 0$ and all sufficiently large index n. The error term $O(\cdot)$ in formula (13b) is uniform in its variables x, y and n.

5.) Let a distribution function F satisfy the conditions formulated in relations (2) and (11). Let us show with the help of formula (13b) the asymptotic formula

$$F_n(\varepsilon\sqrt{n}|y) - F_n(x|y) = (1 - \Phi(x)) \exp\left\{O\left(\frac{x^3 + x^2|y| + |y| + 1}{\sqrt{n}}\right)\right\} \tag{14}$$

for all $n \ge n_0$ with an appropriate threshold number n_0 and a sufficiently small number $\varepsilon > 0$ if $0 \le x \le \frac{\varepsilon}{4}\sqrt{n}$ and $|y| \le \frac{\varepsilon}{4}\sqrt{n}$.

b) Let us give under the above conditions on the function F a good upper bound on $1 - F_n(\varepsilon\sqrt{n})$ with the help of relation (12b) if $n \ge n_0$ with a sufficient number n_0, $0 \le x \le \frac{\varepsilon}{4}\sqrt{n}$ and $|y| \le \frac{\varepsilon}{4}\sqrt{n}$. Prove with the help of this estimate and the formula in (14) that if a distribution function F satisfies both relations (2) and (11), then *Property A* holds for it.

In the next problem it will be shown that there exists a distribution function F satisfying relation (2) but not satisfying *Property A*. This counter example is based on the following idea: Let us take a distribution function F which is concentrated on such points $x_1, x_2, \dots$ which are linearly independent over the rational numbers, i.e. if $\sum\limits_{j=1}^{k} r_j x_j = 0$ with some positive integer k and rational numbers $r_1, \dots, r_k$, then $r_j = 0$ for all numbers $1 \le j \le k$. If we choose a sequence $X_1, \dots, X_{2n}$ of independent random variables with such a distribution function F, then the value of the sum $S_{2n} = \sum\limits_{k=1}^{2n} X_k$ determines the value of the set of random variables $\{X_1, \dots, X_{2n}\}$, only the indices of the random variables in this set remain undetermined. If the set of these random variable contains an extremely large number, then the value of $S_n - \frac{1}{2}S_{2n}$ is very large or very small depending on whether this value is taken by a random variable with an index smaller or larger than n. Hence for an appropriate not too large number y the conditional distribution of the random variable $\sqrt{\frac{2}{n}}(S_n - \frac{1}{2}S_{2n})$ under the condition $S_{2n} = y\sqrt{2n}$ may strongly differ from the standard normal distribution function.

6.) Let $X_1, X_2, \dots, EX_1 = 0, EX_1^2 = 1$ be independent, identically distributed random variables with expectation zero and variance 1 which take some values $x_1, x_2, \dots,$ with probabilities $p_1, p_2, \dots, p_n > 0$, $n = 1, 2, \dots,$ $\sum\limits_{n=1}^{\infty} p_n = 1$. Put $S_n = \sum\limits_{k=1}^{n} X_k$, $n = 1, 2, \dots$. Let us assume that the numbers $x_1, x_2, \dots$ and $p_1, p_2, \dots$ satisfy the following conditions:

   a.) The numbers x_1, x_2, $\dots$ are independent (in algebraic sense) over the field of rational numbers.

   b.) $n < |x_n| < n + 1$ and $B_1 e^{-n} \le p_n \le B_2 e^{-n}$ with some appropriate constants $0 < B_1 < B_2 < \infty$ for all numbers $n = 1, 2, \dots$.

   Let us define a set $A_n = A_n(p, C)$ with some appropriate constants $0 < p < 1$ and $C > 0$. The set A_n consists of such sequences $\{x_{j_1}, \dots, x_{j_{2n-1}}\}$ of length $2n - 1$ whose elements belong to the above introduced numbers $x_1, x_2, \dots$ and satisfy the following properties.

   (i.) $\left|\sum\limits_{s=1}^{2n-1} x_{j_s}\right| < C\sqrt{n}$

   (ii.) Let us consider all permutations $\pi = \pi_{2n-1} = \{\pi(1), \dots, \pi(2n-1)\}$ of the set $\{1, \dots, 2n-1\}$. There exist more than $p\binom{2n-1}{n-1}$ such permutations π for which the inequality $\left|\sum\limits_{s=1}^{n-1} x_{j_{\pi(s)}}\right| < C\sqrt{n}$ holds.

   Let us show that with an appropriate choice of the constants $C > 0$ and $0 < p < 1$ it can be achieved that $P(\{X_1, \dots, X_{2n-1}\} \in A_n) > q$ with some constant $q > 0$ not depending on the number n. Moreover, it can be achieved that this number $q > 0$ be arbitrarily close to the number 1.

   Given a sequence $\{x_{j_1}, \dots, x_{j_{2n-1}}\} \in A_n$ put $y_1 = \sum\limits_{j=1}^{2n-1} x_{j_1}$, $m = m(y_1) = [y_1]$, where $[u]$ denotes the integer part of the number u, and introduce the numbers $M = M(n) = [5m]$ and $y = \frac{y_1 + x_M}{\sqrt{2n}}$. Then

$$P\left(\left|\sqrt{\frac{2}{n}}\left(S_n - \frac{1}{2}S_{2n}\right)\right| > C \left| \frac{S_{2n}}{\sqrt{2n}} = y\right.\right) \ge \frac{q}{2}.$$

   This implies that the distribution function F of the random variables $X_1, X_2, \dots$ does not satisfy *Property A*. Moreover, the probability of the event that the normalized sum $\frac{S_{2n}}{\sqrt{2n}}$ takes such a value y, for which the conditional distribution function $F_n(x|y)$ at the left-hand side of formula (10) satisfies the inequality

$$\sup_{|x|<K} |F_n(x|y) - \Phi(x)| > \alpha > 0$$

with some appropriate constants $K > 0$ and $\alpha > 0$ is greater than $e^{\text{const.}\sqrt{n}}$.

This means that the probability of existence of such "bad values" of y is relatively large, if we compare it with the probability of the event $P(|S_{2n}| \ge \varepsilon n\}$ which is exponentially small. It was natural to compare the probability of these two events, because *Property A* does not supply a good approximation of the conditional distribution function $F(x|y)$ under the condition $S_{2n} = y\sqrt{2n}$ if $|y|\sqrt{2n} \ge \varepsilon n$.

In Problem 6 such a distribution function F supplied the counter example for *Property A* which is not smooth, and it takes large values with relatively large probability. If *Property A* does not hold, then the proof of the *Finite Version of the Approximation Theorem* has to be modified. I briefly sketch how this result can be proved by a slight modification of the construction applied in its proof if the distribution function F satisfies one of the following conditions.

a.) The distribution function F has a representation $F = pG + (1-p)H$ with two distribution functions G and H such $0 < p \le 1$, and the distribution function G has a density function.

b.) The distribution function F is the distribution of a bounded random variable. That is, there exist some numbers $-\infty < A < B < \infty$ such that $F(A) = 0$ and $F(B) = 1$.

Beside this a Problem 8 will be formulated which enables to reduce the proof of the *Finite Version of the Approximation Theorem* to these two special cases when either condition a) or b) is satisfied. (Let me remark that in the counter example considered in Problem 6 neither condition a) nor condition b) is satisfied.)

It is a natural idea that in the case of such independent and identically distributed random variables $X_1, \dots, X_{2^n}$ whose distribution function F does not satisfy Property A, and as a consequence the *Finite Version of the Approximation Theorem* cannot be satisfied with the help of the previous construction we can try to overcome this difficulty by means of an appropriate smoothing of the distribution function F. We may expect that by adding sufficiently small independent normal random variables η_i with expectation zero and appropriately chosen variance to the random variables X_i (which are independent of them) we get a new sequence of independent identically distributed random variables with smooth distribution function, hence they satisfy a slightly modified version of Property A. Then we may try to apply a natural modification of the original construction in the proof of the *Finite Version of the Approximation Theorem* to this new sequence. In such a way we may prove that this new sequence of independent random variables satisfies the *Finite Version of the Approximation Theorem.* If we can do this with the help of Gaussian random variables η_i with sufficiently small variances, then the result we get for the modified sequence implies automatically the *Finite Version of the Approximation Theorem* for the original sequence of independent random variables. I briefly show that this program can be carried out if the distribution function F satisfies the above formulated condition a). (Several technical details of the proof will be omitted.) On the other hand, we can apply this way of proof only if condition a) holds, because we need the contribution of the absolute continuous part of the distribution function F to get sufficiently strong smoothing effect.

Let $X_1, \dots, X_{2^n}$ be a sequence of independent and identically distributed random variables with a distribution function F satisfying formula (2) and condition a.) and take a sequence $\eta_1, \dots, \eta_{2^n}$ of independent normally distributed random variables with expected value zero and variance $\sigma^2 = 2^{-n}$ which is independent also of the original sequence $X_1, \dots, X_{2^n}$. Define the sequence of random variables $\bar{X}_k = X_k + \eta_k$, $k = 1, \dots, 2^n$. The sequence $\bar{X}_1, \dots, \bar{X}_{2^n}$ consists of independent $\bar{F} = \bar{F}^{(n)} = F * G_{0,2^{-n}}$ distributed random variables, where $G_{0,2^{-n}}$ denotes the normal distribution with expectation zero and variance 2^{-n} and is $*$ the convolution operator. If a Wiener process $W(t)$ is given, then the method of proof of the finite version of the *Finite Version of the Approximation Theorem* enables us to construct a sequence $\bar{X}'_1, \dots, \bar{X}'_{2^n}$ with the same distribution as $\bar{X}_1, \dots, \bar{X}_{2^n}$ such that the partial sums $S'_k = \sum\limits_{j=1}^{k} \bar{X}'_j$, $k = 1, \dots, 2^n$, satisfy an appropriate version of formulas (6) and (6a). We get this version by replacing the random broken line $S_n(t, \omega)$ with the random broken line which appears if we write the random variables $\bar{S}'_k$ instead of S'_k in formula $(1')$. Furthermore, I state that this result also implies that a distribution function F satisfying both formula (2) and condition a) also satisfies the *Finite Version of the Approximation Theorem.*

The existence of a sequence $\bar{X}_1, \dots, \bar{X}_{2^n}$ with the properties mentioned in the last paragraph can be proved by means of the halving procedure with the help of the underlying Wiener procedure. To prove that the random variables obtained in such a way satisfy the appropriate version (6) and (6a) we have to show that if we consider instead of the conditional distribution

$$F_n(x|y) = P\left(\sqrt{\frac{2}{n}}\left(S_n - \frac{1}{2}S_{2n}\right) \le x \,\middle|\, \frac{S_{2n}}{\sqrt{2n}} = y\right)$$

the conditional distribution

$$F^{(n)}_{\bar{n}}(x|y) = P\left(\sqrt{\frac{2}{\bar{n}}}\left(\bar{S}_{\bar{n}} - \frac{1}{2}\bar{S}_{2\bar{n}}\right) \le x \,\middle|\, \frac{\bar{S}_{2\bar{n}}}{\sqrt{2\bar{n}}} = y\right), \tag{15}$$

then this new conditional distribution function satisfies an appropriate version of *Property A*. In the definition of the conditional distribution $F^{(n)}_{\bar{n}}(x|y)$ we have introduced a new parameter $\bar{n}$, and considered the partial sums $\bar{S}_{\bar{n}} = \sum\limits_{j=1}^{\bar{n}} \bar{X}_j$, and give an estimate of the closeness of the conditional distribution function $F^{(n)}_{\bar{n}}(x|y)$ and the standard normal distribution function depending on both parameters n and $\bar{n}$. (In the applications we have in mind the parameter n is fixed at the start as we consider a sequence of length 2^n and choose Gaussian random variables with variance 2^{-n}. In the successive application of the halving procedure we have to investigate the conditional distribution functions $F^{(n)}_{\bar{n}}(x|y)$ with different parameters $\bar{n} = 2^{n-l}$, $l = 1, \dots, n$.) Actually it is enough to prove a good result on the asymptotic behavior of the conditional distribution function $F^{(n)}_{\bar{n}}(x|y)$ only in the case $\bar{n} \ge Kn$ with an appropriate (large) number $K > 0$

not depending on the number n. (To understand why such a reduced version of this estimate is sufficient for our purposes we have to remember that the fluctuation of the Wiener process $W(t,\omega)$ and random broken line process $\bar S_n(t,\omega)$ is relatively small in small intervals. A detailed calculation shows that the fluctuations of these processes is sufficiently small for our purposes in intervals of length $[0,Kn]$. I formulate the version of Property A we need in this case.

The definition of the modified version of Property A. *We say that a sequence of independent identically distributed random variables $X_1,X_2,\dots$, with distribution function F satisfies the modified version of Property A, if the conditional distribution functions $F_{\bar n}^{(n)}(x|y)$ defined in formula (15) satisfy the following asymptotic relation. There exists some numbers $\varepsilon>0$, $K>0$ and threshold index n_0 such that*

$$
\begin{aligned}
1-F_{\bar n}^{(n)}(x|y)&=(1-\Phi(x))\exp\left\{O\left(\frac{x^3+x^2|y|+|y|+1}{\sqrt{\bar n}}\right)\right\}\\
&\qquad \textit{if } \bar n\ge Kn,\ 0\le x\le \varepsilon\sqrt{\bar n},\ \ 0\le |y|\le\varepsilon\sqrt{\bar n}\\
F_{\bar n}^{(n)}(-x|y)&=(1-\Phi(x))\exp\left\{O\left(\frac{x^3+x^2|y|+|y|+1}{\sqrt{\bar n}}\right)\right\}\\
&\qquad \textit{if } \bar n\ge Kn,\ 0\le x\le \varepsilon\sqrt{\bar n},\ \ 0\le |y|\le\varepsilon\sqrt{\bar n},
\end{aligned}
\tag{10a}
$$

holds, where $\Phi(x)$ is the standard normal distribution function, and the error term $O(\cdot)$ is uniform in the variables x, y, n and $\bar n$.

To prove the *Modified version of Property A* if the distribution function F satisfies relation (2) and condition a) in the same way as the original *Property A* was proved it is enough to show that the density functions $f_{\bar n}^{(n)}(x)$ of the normalized partial sums $\frac{\bar S_{\bar n}}{\sqrt{\bar n}}=\frac{1}{\sqrt{\bar n}}\sum\limits_{j=1}^{\bar n}\bar X_j$ satisfy such a version of relations (12a) and (12b) in the *Sharp form of the local central limit theorem* where the number n is replaced by $\bar n$ everywhere at the right-hand side of these relations, and it is assumed that $\bar n\ge Kn$.

This version of the *Sharp form of the Local Central Limit Theorem* can be proved by the method of the solution of Problem 23 in the *Theory of Large Deviations I.* if the distribution function $F(x)$ satisfies condition a.). The main idea of the proof is that the density function we want to estimate can be expressed by the inverse Fourier transform of the characteristic function, or by the analytic continuation of this formula, provided that the characteristic function and its analytic continuation is an integrable function. Beside this, the expression we get in such a way can be well investigated.

We have to study the expressions in the following identity:

$$
\sqrt{\bar n}f^{(n)}(\sqrt{\bar n}x)=\frac{1}{2\pi}\int e^{(is-t)x}\frac{\bar R_{\bar n}(s+it)}{\bar R_{\bar n}(it)}\,ds,
$$

where $\bar R_{\bar n}(s+it)=\left(R(s+it)e^{2^{-n-1}(t^2-s^2)}\right)^{\bar n}$, and $R(s+it)=\int e^{(is-t)x}F(\,dx)$ is the analytic continuation of the characteristic function of the $F(x)$ distribution function.

This means that $\bar R_{\bar n}(s+it)$ is the analytic continuation of the distribution function $F^{(n)}_{\bar n}(x)$. This function, as the function of the variable s with a fixed parameter t is integrable, since the function $e^{2^{-n-1}(t^2-s^2)}$ is integrable, and the function $R(s+it)$ is bounded. But in the proof of the modified version of the *Sharp form of the Local Central Limit Theorem* we need some more information. We have to know that the integral expressing the density function as the inverse Fourier transform of the characteristic function and its analytic continuation is essentially localized in a small neighbourhood of the origin, where the integrand can be well estimated. Condition a) was imposed to guarantee this property. The consequence of condition a) needed for us is formulated in the following Problem 7.

7.) If the distribution function F satisfies condition a), then for all numbers $A>0$ and $B>0$ there exists some number $\alpha=\alpha(A,B)<1$ such that

$$\left|\frac{R(s+it)}{R(it)}\right| < \alpha \quad \text{if } |s|>A \quad \text{and} \quad |t|<B,$$

where $R(s+it)=\int e^{(is-t)x}F(\,dx)$.

The result of Problem 7 together with the fact that the function $e^{2^{-n-1}\bar n(t^2-s^2)}$ (as a function of the variable s with a fixed t) is integrable, and the integral of this function is not too large, guarantees that the localization property we need in the proof of the modified version of the *Sharp form of the Local Central Limit Theorem* can be proved if condition a) holds, and $\bar n\ge Kn$. Here I omit the discussion of the technical details.

The Finite version of the Approximation Theorem also holds if condition b) holds, but in this case we can prove this statement with the help of a modified version of the construction and with different justification of this method.

In this case we can apply the following modified version of the halving procedure. We have a Wiener process $W(t,\omega)$, $0\le t\le 2^n$, at the start. Step zero of our procedure is carried out in the usual way; the random sum $S_{2^n}(\omega)$ is defined as the quantile transform of the random variable $W(2^n,\omega)$, of the value of the Wiener process $W(t,\omega)$ in its end-point $t=2^n$. After this we construct in the knowledge of the value of the random sum $S_{2^n}(\omega)$ the set of random variables $\{X_1(\omega),\dots,X_{2^n}(\omega)\}$ with the right conditional distribution in such a way that their sum equal $S_{2^n}(\omega)$. In the construction of this set we apply beside the random variable $S_{2^n}(\omega)$ such random variables which are independent of the Wiener process $W(t,\omega)$. At this step we define the value of all random variables $\{X_1(\omega),\dots,X_{2^n}(\omega)\}$, but do not tell their indices. Let us observe that the conditional distribution of all possible indexations of this set under the condition that the set of values of our random variables is prescribed, only the index of the random variable which takes a given value is not known has the same probability $(2^n!)^{-1}$. If we define the indexation in such a way that the probability of all possible indexation equals $(2^n!)^{-1}$ in the case of all possible set of values $\{X_1(\omega),\dots,X_{2^n}(\omega)\}$, then the random variables $X_1(\omega),\dots,X_{2^n}(\omega)$ constructed in such a way are independent with distribution F. On the other hand, we want to make this random indexation in such a way that the random sums $S_k(\omega)=\sum\limits_{j=1}^{k} X_k(\omega)$ be close to $W(k,\omega)$, to the value of the Wiener process $W(t,\omega)$ in the point $t=k$ for all values $1\le k\le 2^n$ $W(k,\omega)$.

We define the indexation by an inductive procedure. In the first step of this procedure we tell with the help of the random variable $2W(2^{n-1},\omega)-W(2^n,\omega)$ which elements of the set $\{X_1(\omega),\dots,X_{2^n}(\omega)\}$ have an index less than or equal to 2^{n-1}, and which elements have an index greater than 2^{n-1}. We want to do this in such a way that all subsets of the set $\{1,2,\dots,2^n\}$ with 2^{n-1} elements is chosen with the same probability for the set of indices of the (random) subset $\{X_1(\omega),\dots,X_{2^{n-1}}(\omega)\}$. Let us observe that by defining this set we also define the value of the random variable $2S_{2^{n-1}}(\omega)-S_{2^n}(\omega)$. We want to make the first step of the construction in such a way (with the help of the quantile transformation) that the difference $\left[2W(2^{n-1},\omega)-W(2^n,\omega)\right]-\left[2S_{2^{n-1}}(\omega)-S_{2^n}(\omega)\right]$ be small. We try to make a similar construction also in the subsequent steps of the procedure.

After the l-th step of our construction we have determined the random sets

$$\{X_{k2^{n-l}+1}(\omega),\dots,X_{(k+1)2^{n-l}}(\omega)\},\quad 0\le k\le 2^l-1,$$

but we do not know the indices of the individual random variables in this set. In the $l+1$-th step we tell with the help of the random variable

$$[W((2k+1)2^{n-l-1},\omega)-W(k2^{n-l}+1,\omega)]-[W((k+1)2^{n-l},\omega)-W((2k+1)2^{n-l-1},\omega)]$$

which elements of this set have an index less than or equal to $(2k+1)2^{n-l-1}$. We choose this random set of indices in such a way that all subsets of $\{k2^{n-l}+1,k2^{n-l}+2,\dots,(k+1)2^{n-l}\}$ of 2^{n-l-1} elements are chosen with the same probability for this set. Beside this, we make this halving of the sets $\{k2^{n-l}+1,k2^{n-l}+2,\dots,(k+1)2^{n-l}\}$ for different indices k, $0\le k<2^l$, independently of each other. We also want to achieve (with the application of the conditional quantile transform) that the random variables

$$[W((2k+1)2^{n-l-1},\omega)-W(k2^{n-l}+1,\omega)]-[W((k+1)2^{n-l},\omega)-W((2k+1)2^{n-l-1},\omega)] \tag{16a}$$

and

$$[S_{(2k+1)2^{n-l-1}}(\omega)-S_{k2^{n-l}+1}(\omega)]-[S_{(k+1)2^{n-l}}(\omega)-S_{(2k+1)(2^{n-l-1})}(\omega)] \tag{16b}$$

be close to each other.

We make the $l+1$-th step of the halving procedure by defining first the random variables in (16b) by calculating the distributions of the expressions in (16b) (which depends on the elements in the k-th block, hence on the number of index k, and then by constructing the random variables in (16b) by means of the quantile transform from the random variables in (16a). In such a way we prescribe the value of the random sums $S_{(2k+1)2^{n-l-1}}(\omega)-S_{k2^{n-l}+1}(\omega)$ for all $k=1,\dots,2^l$. If it determines the value of the terms taking part in this sum in a unique way, then the indices of the terms in this sum constitute the set $\{k2^{n-l}+1,\dots,(2k+1)2^{n-2},\}$. If there are several possibilities for writing down this random variable as the sum of 2^{n-l-1} terms of the prescribed numbers, then we choose one of them randomly, by choosing all possibilities with equal probability, and the indices of these terms will belong to the set $\{k2^{n-l}+1,\dots,(2k+1)2^{n-l-1},\}$. Let us also observe that the random variables in (16a) are independent for different indices l or k. This fact guarantees the independence we need in the halving procedure.

Let us remark that in the case when the distribution function F of the random variables $X_k(\omega)$, $1\le k\le 2^n$, is concentrated in a set of numbers linearly independent over the set of rational numbers (such a case was considered in the counter example of Problem 6) then the previously described construction agrees with the original construction in the proof of *The finite version of the approximation theorem*. We want to show that if the distribution function F satisfies condition b), then the above construction satisfies the desired estimate. This means that if the random variables with distribution F are bounded, then a situation similar to the counter example of Problem 6 cannot appear. To prove that the construction described above yields the above result we need the following theorem.

Theorem B. *Let $2N$ real numbers $x_1,\dots,x_{2N}$ be given which satisfy the condition*

$$\max_{1\le k\le 2N}|x_k|\le K, \qquad \textit{and} \quad \sigma^2=\sum_{k=1}^{2N}(x_k-\bar x)^2\ge cN, \quad \textit{if} \quad \bar x=\frac{1}{2N}\sum_{k=1}^{2N}x_k$$

with appropriate constants $K>0$ and $c>0$. Let us choose randomly one of the permutations $\{\pi(1),\dots,\pi(2N)\}$ of the numbers $1,\dots,2N$, by choosing all possible permutations with probability $\frac{1}{(2N)!}$, and define the random variable

$$S_N=\left(x_{\pi(1)}+\cdots+x_{\pi(N)}\right)-\left(x_{\pi(N+1)}+\cdots+x_{\pi(2N)}\right)$$

It satisfies the following form of the central limit theorem and its large deviation version:

$$P\left(S_N>\sigma x\sqrt N\right)=\left(1-\Phi\left(\frac{x}{\sqrt N}\right)\right)\exp\left\{O\left(\frac{x^3+1}{\sqrt N}\right)\right\},$$
$$P\left(S_N<-\sigma x\sqrt N\right)=\Phi\left(-\frac{x}{\sqrt N}\right)\exp\left\{O\left(\frac{x^3+1}{\sqrt N}\right)\right\}$$

for all numbers $0\le\varepsilon\sqrt N$ with some appropriate number $\varepsilon=\varepsilon(c,K)>0$, where the error term $O(\cdot)$ means the absolute value of the difference of the left-hand side and the main term at the right-hand side is less than $B\frac{x^3+1}{\sqrt N}$ with a constant B depending only on the parameters C and K, but not on the numbers x and N.

The proof of the (non-trivial) Theorem B will be omitted. It can be found in the proof of Lemma 3 of the work of János Komlós, Péter Major and Gábor Tusnády *An approximation of Partial Sums of Independent RV'-s and the Sample DF. II.* Zeitschrift für Wahrscheinlichkeitstheorie **34** (1976) 34–58. Here I only present a heuristic explanation of this result. I also omit the details of the proof of the *Finite Version of the Approximation Theorem* in the case when condition b) holds.

Let us make a random pairing $(x_{j_{2k}}, x_{j_{2k+1}})$, $1 \le k \le N$, of the numbers $x_1, \dots, x_{2N}$, define independent, identically distributed random variables $r_1, \dots, r_N$ such that $P(r_k = 1) = P(r_k = -1) = \frac{1}{2}$, $1 \le k \le N$, and introduce the random variable $U = \sum\limits_{k=1}^{N} r_k \left(x_{j_{2k}} - x_{j_{2k+1}}\right)$. The random variable U is the sum of independent random variables with expectation zero, hence we can estimate its distribution well by means of a normal distribution function with expectation zero and appropriate variance. But this variance depends on the pairing $(x_{j_{2k}}, x_{j_{2k+1}})$, $1 \le k \le N$, of the numbers we consider. The distribution of the random variable S_N considered in Theorem B equals the average of the distributions of the (almost normal) random variables U corresponding to all possible pairings of the numbers $x_1, \dots, x_{2N}$. To prove that this average satisfies the statement of Theorem B it is enough to show that the variances of the distributions taking part in this average are typically very close to the number $N\sigma^2$. The proof of this non-trivial statement is the most important step of the proof.

In the next step I formulate Problem 8 which enables us to reduce the proof of the *Finite Version of the Approximation Theorem* to the two special cases when the distribution function F of the independent random variables we are investigating satisfies either condition a) or condition b).

8.) Let us fix some distribution functions F_1, F_2 and G_1, G_2. Let S_n^i, and T_n^i, $n = 1, 2, \dots,$ be the sequences of partial sums of independent, identically distributed random variables with distribution functions F_i and G_i, $i = 1, 2$. Let us fix some number $0 \le p \le 1$ and define the distribution functions $F = pF_1 + (1-p)F_2$ and $G = pG_1 + (1-p)G_2$. Let us show that some pairs S_n and T_n, $n = 1, 2, \dots,$ of sequences of partial sums of independent, identically distributed random variables can be constructed with distribution functions F and G in such a way that they satisfy the relation

$$
\begin{aligned}
&P\left(\sup_{1 \le j \le n} |S_j - T_j| \ge a + b\right) \\
&\qquad \le P\left(\sup_{1 \le j \le n} \left|S_j^{(1)} - T_j^{(1)}\right| \ge a\right) + P\left(\sup_{1 \le j \le n} \left|S_j^{(2)} - T_j^{(2)}\right| \ge b\right)
\end{aligned}
$$

for arbitrary real numbers $a > 0$, $b > 0$ and integer $n > 0$.

Let us reduce with the help of the above statement the proof of the *Finite Version of the Approximation Theorem* to the two special cases when the distribution function F of the independent random variables we are investigating satisfy one of the conditions a) or b).

In the next problems we investigate the converse of the above Approximation Theorem, that is we are interested in the question which are the lower bounds for the possibility of approximation of partial sums with Wiener process or of normalized the empirical distribution function by Brownian bridge. The proof of these lower bounds is based on such estimates which give a lower bound on the possibility of approximation of the distribution function of partial sums of independent random variables by means of a normal distribution function. These estimates belong to the estimates of the theory of the central limit theorem and large deviation theory. Because of some technical reasons it is more convenient to work with the moment generating functions of our random variables instead of their distribution. The result of the next Problem 9 has such a content.

9.) Let $F(x)$ be such a distribution function for which the moment generating function $R(s) = \int e^{sx} F(\,dx)$ exists in some interval $-a < s < a$, $a > 0$. The value of the moment generating function $R(s)$ in the interval $[-a, a]$ uniquely determines the distribution function $F(x)$. (This number $a > 0$ can be chosen sufficiently small.)

In the next problem we prove formula (5) in the case when the random variable X has moment generating function in a small neighbourhood of the origin.

10.) Let $X_1, X_2, \dots,$ be a sequence of independent, identically distributed random variables which satisfy the relation $R(2s) = Ee^{2sX_1} < \infty$ with some number $s > 0$. Let us fix some positive integer n and define the random variables $S_{k,n} = \sum\limits_{j=kn+1}^{k(n+1)} X_j$, $k = 1, 2, \dots$. Let us choose a sufficiently large number $A > 0$ and put $N(n) = e^{An}$. Then the relation

$$\lim_{n\to\infty} \frac{1}{N(n)R^n(s)} \sum_{k=1}^{N(n)} e^{sS_{k,n}} = 1 \quad \text{with probability 1} \tag{17}$$

holds if $A > 0$ is sufficiently large ($N(n) = e^{An}$), and $R(s) = Ee^{sX_1}$.

Let $Y_1, Y_2, \dots,$ be a sequence of independent random variables with standard normal distribution and put $T_{k,n} = \sum\limits_{j=kn+1}^{k(n+1)}$ with some appropriate real number n. Let us observe that such a version of relation (17) holds in which the random variable $S_{k,n}$ is replaced by $T_{k,n}$ and the moment generating function $R(s)$ by $\bar{R}(s) = Ee^{sY_1} = e^{s^2/2}$. Furthermore by the result of the previous problem there exists an arbitrary small number $s > 0$ for which $R(s) \neq \bar{R}(s)$ if X_1 is not a standard normal random variable. Let us prove with the help of the above observation formula (5) if the random variable X_1 has moment generating function formula (5) in a small neighbourhood of the origin.

The result of the next problem is about approximation of the normalized empirical distribution function by a Brownian bridge, and it is the analog of the previous result.

11.) Let $Z_n(t)$, $0 \le t \le 1$, be a normalized empirical distribution with n sample points. (This means that there are n independent random variables $\xi_1, \dots, \xi_n$ with uniform distribution on the interval $[0,1]$, and we consider the random process $Z_n(t) = \frac{1}{\sqrt{n}}(P_n(t)-nt)$, where $P_n(t) = \sum\limits_{j=1}^{n} I(\xi_j < t)$, and $I(A)$ denotes the indicator function of the set A.) Beside this, let $X_n(t)$, $0 \le t \le 1$, be a Brownian bridge on the same probability space where the random process $Z_n(t)$ is defined. (The distribution of the process $X_n(t)$ does not depend on the number n.) Let us fix a sufficiently small number $c > 0$, and define the numbers $u_k = k\frac{c\log n}{n}$, $0 \le k \le M(n)$, where $M(n) = \left[\frac{n}{c\log n}\right]$, and $[\cdot]$ denotes integer part. Let us fix the random variables $U_k = \sqrt{n}\,(X(u_k) - X(u_{k-1}))$ and $V_k = U_{k,n} = \sqrt{n}\,(Z_n(u_k) - Z_n(u_{k-1}))$, $1 \le k \le M(n)$ Fix a number $t > 0$ and prove the following relations:

$$\frac{1}{M(n)\bar{R}(n)} \sum_{k=1}^{M(n)} e^{tV_k} \Rightarrow 1$$

$$\frac{1}{M(n)R(n)} \sum_{k=1}^{M(n)} e^{tU_k} \Rightarrow 1$$

where the number $c > 0$ is sufficiently small, and $\Rightarrow$ denotes stochastic convergence,

$$\bar{R}(n) = Ee^{tV_1} = \exp\left\{c\log n(e^t - 1 - t) + O\left(\frac{(\log n)^2}{n}\right)\right\},$$

and

$$R(n) = Ee^{tU_1} = \exp\left\{\frac{1}{2}c\log n\left(1 - \frac{c\log n}{n}\right)\right\}.$$

Let us show with the help of the above statements that there exists a sufficiently small number $K > 0$ such that

$$P\left(\sqrt{n}\sup_{0\le t\le 1}(Z_n(t) - X_n(t)) > K\log n\right) \to 1 \quad \text{if } n \to \infty.$$

In the next problem we consider the sequence of independent, identically distributed random variables and the partial sums made from them in the case when some moment type function of these random variables in infinite. We shall show that in this case the fluctuation between the neighbouring terms of the partial sums is sometimes very large, and this yields a lower bound for the approximation of these partial sums by the partial sums of independent standard normal random variables. A Special case of this result is *Statement 2* which also contains that part of *Statement 1* not considered in Problem 10 which deals with the case when the random variables we consider have no distribution function.

12.) Let $X_1, X_2, \dots,$ be a sequence of independent and uniformly distributed random variables, and put $S_n = \sum\limits_{k=1}^{n} X_k$, $n = 1, 2, \dots$, $X_k^+ = \max(X_k, 0)$, $k = 1, 2, \dots$. Let $H(x)$ be a continuous, strictly monotone function defined on the set of non-negative real numbers for which $H(0) = 0$, $\lim\limits_{x\to\infty} H(x) = \infty$, the numbers K_n, $n = 1, 2, \dots$ as the solution of the equation $H(x) = n$. Then

$$S_n - S_{n-1} \ge K_n \quad \text{with probability 1 for infinitely many indices } n$$
$$\text{if and only if } EH(X_1^+) = \infty$$

Let $EH(X_1^+) = \infty$, and let us also assume that $H(x) \le e^{\alpha x}$ with some number $\alpha > 0$. Let $Y_1, Y_2, \dots$ be a sequence of independent random variables with standard normal distribution function, and put $T_n = \sum\limits_{k=1}^{n} X_k$, $n = 1, 2, \dots$. Then

$$\limsup_{n\to\infty} \frac{|S_n - T_n|}{2K_n} \ge 1 \quad \text{with probability 1.} \tag{18}$$

In particular, if $Ee^{\alpha X_1^+} = \infty$ or $Ee^{\alpha X_1^-} = \infty$ with some number $\alpha > 0$, where $X_1^- = -\min(X_1, 0)$, then relation (18) holds with the choice $K_n = \frac{1}{\alpha}\log n$. If $E|X_1|^r = \infty$ with some number $r > 0$, then relation (18) holds with the choice $K_n = Cn^{1/r}$, where $C > 0$ can be arbitrarily large fixed positive number.

In the last two problems such results are considered which may be useful in an overview of this subject.

13.) Let X_n, $n = 1, 2, \dots,$ be a sequence of independent random variables with normal distribution such that $EX_n = 0$, $EX_n^2 = \sigma_n^2$, $0 < \sigma_n < 1$, $n = 1, 2, \dots,$ and $\lim\limits_{n\to\infty} \sigma_n^2 = 1$. Let us introduce the sequence of the partial sums $S_n = \sum\limits_{k=1}^{n} X_k$, $n = 1, 2, \dots$. There exists such a sequence Y_n, $n = 1, 2, \dots,$ (for instance the choice $Y_n = \frac{X_n}{\sigma_n}$, $n = 1, 2, \dots,$ is an appropriate choice) for which the partial sums $T_n = \sum\limits_{k=1}^{n} Y_k$, $n = 1, 2, \dots,$ satisfy the relation

$$\lim_{n\to\infty} \frac{|S_n - T_n|}{\sqrt{n \log\log n}} = 0 \quad \text{with probability 1.}$$

On the other hand this relation is sharp. To formulate this statement more explicitly let us define the numbers $D_n^2 = \sum\limits_{k=2^{n-1}+1}^{2^n} \sigma_k^2$. Let us show that for an arbitrary sequence u_n, $u_n \ge 1$, $n = 1, 2, \dots,$ the event

$$S_{2^n} - S_{2^{n-1}} \ge D_n u_n$$

holds with probability 1 for infinitely or finitely many indices n depending on the convergence or divergence of the sum $\sum\limits_{n=1}^{\infty} \frac{e^{-u_n^2/2}}{u_n}$. Let us show with the help of this result that for every sequence $f(n)$, $f(n) > 0$, $n = 1, 2, \dots$, such that $\lim\limits_{n\to\infty} f(n) = \infty$, there exists a sequence σ_n, $0 < \sigma_n \le 1$, $\lim\limits_{n\to\infty} \sigma_n = 1$, $n = 1, 2, \dots$, such that if X_n, $n = 1, 2, \dots$, is a sequence of independent Gaussian random variables with expectation zero and variance $EX_n^2 = \sigma_n^2$, Y_n, $n = 1, 2, \dots$, is a sequence of independent standard normal random variables, and S_n and T_n, $n = 1, 2, \dots$, denotes the partial sums of these random variables, then they satisfy the relation

$$\limsup_{n\to\infty} f(n) \frac{|S_n - T_n|}{\sqrt{n \log\log n}} = \infty \quad \text{with probability 1.}$$

14.) Let four urns be given. Let us throw M balls to these urns independently of each other in such a way that every ball falls with the same probability $\frac{1}{4}$ in each urn. Let $X_j = X_j(M)$, $j = 1, 2, 3, 4$, denote the number of balls falling in the j-th urn. Let us prove the identity which tells the conditional probability of the event that a prescribed number of balls fall into the first urn, under the condition that the number of balls falling in the first and second urn and the number of the balls falling in the first or third urn is prescribed. (This condition can be rewritten in an equivalent form by prescribing the number of balls falling in the third or fourth urns and the number of balls falling in the second and fourth urns.) Let us prove the following identity:

$$P(X_1 = k | X_1 + X_2 = U, X_1 + X_3 = V) = \frac{\binom{U}{k}\binom{M-U}{V-k}}{\binom{M}{V}}$$

under the condition that $0 \le U \le M$, $0 \le V \le M$, $0 \le k \le \min(U, V)$.

Solutions.

4. To prove relation (13a) let us first make the following observation. If (X, Y) are two random variables whose joint distribution has a density function $g(x, y)$, then the conditional distribution $G(x|y) = P(2^{-1/2}(Y - X) < x | 2^{-1/2}(Y + X) = y)$ has a density function for all parameters y, and it equals $g(x|y) = \frac{g\left(\frac{y-x}{\sqrt{2}}, \frac{y+x}{\sqrt{2}}\right)}{\sqrt{2}h\left(\sqrt{2}y\right)}$, where $h(y) = \int_{-\infty}^{\infty} h(u, y - u)\, du$ is the density function of the random variable $X + Y$. We get formula (13a) by applying this relation with the choice $X = n^{-1/2}S_n$, $Y = n^{-1/2}(S_{2n} - S_n)$. Indeed, since X and Y are independent random variables with density function $f_n(x)$, hence their joint density exists, and it equals $g(x, y) = f_n(x)f_n(y)$. Beside this, $h_n(y) = \frac{1}{\sqrt{2}} f_{2n}\left(\frac{y}{\sqrt{2}}\right)$. (Let us remark that if the k-th power of the Fourier transform is integrable, then the partial sums S_n have a density function for $n \ge k$.)

Formulas (12a) and (13a) imply that

$$
\begin{aligned}
f_n(x|y) = \frac{\varphi\left(\frac{y+x}{\sqrt{2}}\right)\varphi\left(\frac{y-x}{\sqrt{2}}\right)}{\varphi(y)} \exp\Biggl\{ n\Biggl(\left(\frac{x+y}{\sqrt{2n}}\right)^3 \lambda\left(\frac{x+y}{\sqrt{2n}}\right) \\
+ \left(\frac{y-x}{\sqrt{2n}}\right)^3 \lambda\left(\frac{x-y}{\sqrt{2n}}\right) - 2\left(\frac{y}{\sqrt{2n}}\right)^3 \lambda\left(\frac{y}{\sqrt{2n}}\right)\Biggr)\Biggr\} \\
\exp\left\{O\left(\frac{1+|x|+|y|}{\sqrt{n}}\right)\right\}
\end{aligned}
\tag{A1}
$$

if $|x| \le \varepsilon\sqrt{n}$ and $|y| \le \varepsilon\sqrt{n}$ with a sufficiently small $\varepsilon > 0$. To get a good asymptotic on the right-hand side of formula (A1) introduce the function $\bar\lambda(u) = u^3\lambda(u)$. This function is, together with the function $\lambda(u)$, analytic in a small neighbourhood of zero. Hence a Taylor expansion around the point v yields that $\bar\lambda(v+u) + \bar\lambda(v-u) - 2\bar\lambda(u) = \bar\lambda''(v)u^2 + O(u^4) = O(|v|u^2 + u^4)$ if $|u| \le \varepsilon$, $|v| \le \varepsilon$ with some sufficiently small $\varepsilon > 0$, and the $O(\cdot)$ is uniform in both variables u and v. At the end of the above estimate we have exploited that $|\bar\lambda''(v)| \le \text{const.}\,|v|$ in a small neighbourhood of zero. By applying the above formula with $u = \frac{x}{\sqrt{2n}}$ and $v = \frac{u}{\sqrt{2n}}$ we get that

$$
\begin{aligned}
n\left(\left(\frac{x+y}{\sqrt{2n}}\right)^3 \lambda\left(\frac{x+y}{\sqrt{2n}}\right) + \left(\frac{y-x}{\sqrt{2n}}\right)^3 \lambda\left(\frac{x-y}{\sqrt{2n}}\right) - 2\left(\frac{y}{\sqrt{2n}}\right)^3 \lambda\left(\frac{y}{\sqrt{2n}}\right)\right) \\
= O\left(\frac{x^2|y|}{\sqrt{n}} + \frac{x^4}{n}\right) = O\left(\frac{x^2|y| + |x|^3}{\sqrt{n}}\right)
\end{aligned}
$$

if $|x| \le \varepsilon\sqrt{n}$ and $|y| \le \varepsilon\sqrt{n}$.

Hence relation (A1) yields that

$$f_n(x|y) = \varphi(x)\exp\left\{O\left(\frac{x^2|y|+|x|^3+1+|x|+|y|}{\sqrt{n}}\right)\right\}$$
$$= \varphi(x)\exp\left\{O\left(\frac{x^2|y|+|x|^3+1+|y|}{\sqrt{n}}\right)\right\}$$

if $|x| \le \varepsilon\sqrt{n}$, $|y| \le \varepsilon\sqrt{n}$ with a sufficiently small $\varepsilon > 0$, as we claimed.

5.) Because of formula (13b) there exists such constant $K > 0$ such that

$$\begin{aligned} &\left(1-\frac{K(|y|+1)}{\sqrt{n}}\right)\int_x^{\varepsilon\sqrt{n}}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{s^2}{2}\left(1+K\frac{s+|y|}{\sqrt{n}}\right)\right\}ds \\ &\qquad \le F_n(\sqrt{n}\varepsilon|y)-F_n(x|y)=\int_x^{\sqrt{n}\varepsilon}f_n(s|y)\,ds \qquad \text{(A2)} \\ &\qquad\qquad \le \left(1+\frac{K(|y|+1)}{\sqrt{n}}\right)\int_x^{\varepsilon\sqrt{n}}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{s^2}{2}\left(1-K\frac{s+|y|}{\sqrt{n}}\right)\right\}ds \end{aligned}$$

if $0 \le x \le \frac{\varepsilon}{4}\sqrt{n}$, $|y| \le \frac{\varepsilon}{4}\sqrt{n}$.

To get a good estimate for the upper and lower bound in formula (A2) let us take the change of variable $v^2(s) = s^2\left(1+K\frac{s+|y|}{\sqrt{n}}\right)$ to bound the integral in the lower bound and the change of variable $u^2(s) = s^2\left(1-K\frac{s+|y|}{\sqrt{n}}\right)$ to bound the integral in the upper bound. In the next calculations I show that $\dfrac{ds}{du}-1$ and $\dfrac{ds}{dv}-1$ are very small, hence their contribution to the integrals we get after the change of variables can be considered as part of the error term. In the following calculations I shall assume that $n \ge n_0$ with some appropriate threshold index n_0. In this case all steps we shall do is legitime.

The inequality $u(s) \le s$ holds, beside this $u(s) \ge \frac{3}{4}s$ if $\varepsilon \ge 0$ is sufficiently small, and $x \le s \le \sqrt{n}\varepsilon$. Hence we can write $\frac{ds}{du} = \frac{u}{s\left(1-\frac{K|y|}{\sqrt{n}}-\frac{3Ks}{2\sqrt{n}}\right)} \le \frac{1}{\left(1-\frac{K|y|}{\sqrt{n}}-\frac{3Ks}{2\sqrt{n}}\right)} \le 1+\frac{4K}{\sqrt{n}}(u+|y|)$. Similarly, we can write $\frac{ds}{dv} \ge 1-\frac{4K}{\sqrt{n}}(v+|y|)$, $x \le s \le \varepsilon\sqrt{n}$. Thus we get from relation (A2) that

$$\begin{aligned} &\left(1-\frac{K(|y|+1)}{\sqrt{n}}\right)\int_{v(x)}^{\varepsilon\sqrt{n}}\frac{1}{\sqrt{2\pi}}e^{-v^2/2}\left(1-\frac{4K}{\sqrt{n}}(v+|y|)\right)dv \\ &\qquad \le F_n(\sqrt{n}\varepsilon|y)-F_n(x|y) \qquad \text{(A3)} \\ &\qquad\qquad \le \left(1+\frac{K(|y|+1)}{\sqrt{n}}\right)\int_{u(x)}^{\varepsilon\sqrt{n}}\frac{1}{\sqrt{2\pi}}e^{-u^2/2}\left(1+\frac{4K}{\sqrt{n}}(u+|y|)\right)du \end{aligned}$$

with $v(x) = x\left(1+\frac{K}{\sqrt{n}}(x+|y|)\right)^{1/2}$ and $u(x) = x\left(1-\frac{K}{\sqrt{n}}(x+|y|)\right)^{1/2}$. (Since $v(x) \ge x$ and $u(x) \le x$, we have decreased the lower bound and increased the upper bound in (A3) by writing $\varepsilon\sqrt{n}$ as the upper bound in the integral, and this is allowed.)

To estimate the expressions in formula (A3) let us observe that the primitive function of $ue^{-u^2/2}$ is $-e^{-u^2/2}$, and the standard normal distribution and density functions can be well compared. The following calculation is useful for us if we want to rewrite formula (A3) in a form more appropriate for us: $K(x+1)[\Phi(v(x)-\Phi(\varepsilon\sqrt{n})] \le \varphi(v(x)-\varphi(\varepsilon\sqrt{n}) \le \varphi(u(x)-\varphi(\varepsilon\sqrt{n}) \le K(x+1)[\Phi(u(x)-\Phi(\varepsilon\sqrt{n})]$ with some appropriate $K>0$, where $\varphi(\cdot)$ denotes the standard normal density and $\Phi(\cdot)$ the standard normal distribution function. In the proof of this relation we can exploit that $\frac{x}{2} \le v(x) \le u(x) \le 2x \le \frac{\varepsilon}{2}\sqrt{n}$. In particular, the contribution of the term $\Phi(\varepsilon\sqrt{n})$ is negligible in the above estimates. Some calculation with the help of the above bound and relation (A3) yield the following estimate. There exists some constant $\bar{K}>0$ depending only on the distribution function F such that

$$\begin{aligned}\left(1-\frac{\bar{K}(|y|+x+1)}{\sqrt{n}}\right)[\Phi(\varepsilon\sqrt{n})-\Phi(v(x))] &\le F_n(\sqrt{n}\varepsilon|y)-F_n(x|y)\\ \le \left(1+\frac{\bar{K}(|y|+x+1)}{\sqrt{n}}\right)&[\Phi(\varepsilon\sqrt{n})-\Phi(u(x))]\end{aligned} \tag{A4}$$

with the above defined functions $v(x)$ and $u(x)$ if $0 \le x \le \frac{\varepsilon}{4}\sqrt{n}$ and $|y| \le \frac{\varepsilon}{4}\sqrt{n}$. Moreover, the inequality

$$\begin{aligned}\left(1-\frac{\bar{K}(|y|+x+1)}{\sqrt{n}}\right)[1-\Phi(v(x))] &\le F_n(\sqrt{n}\varepsilon|y)-F_n(x|y)\\ \le \left(1+\frac{\bar{K}(|y|+x+1)}{\sqrt{n}}\right)&[1-\Phi(u(x))]\end{aligned} \tag{A4$'$}$$

holds with possibly different constant $\bar{K}>0$. To see this, it is enough to observe that $1-\Phi(\varepsilon\sqrt{n})$ is much smaller $1-\Phi(v(x))$. Hence the increase we commit by writing $1-\Phi(v(x))$ instead of $\Phi(\varepsilon\sqrt{n})-\Phi(v(x))$ can be compensated by writing a larger constant $\bar{K}$ at the left-hand side of (A4). It is enough to observe that $1-\Phi(\varepsilon\sqrt{n}) \le \frac{1}{\sqrt{n}}[1-\Phi((v(x))]$, since $v(x) \le \frac{\varepsilon}{2}\sqrt{n}$. The replacement of $\Phi(\varepsilon\sqrt{n}))-\Phi(u(x))$ by $1-\Phi(u(x))$ at the right-hand side of (A4) is clearly allowed.

To prove formula (14) with the help of relation (A4$'$) we have to compare $1-\Phi(v(x))$ and $1-\Phi(u(x))$ with $1-\Phi(x)$. We can write, by exploiting that the derivative of $\log[1-\Phi(x)]$ is $\frac{-\varphi(x)}{1-\Phi(x)}$ which can be well bounded that

$$\begin{aligned}0 \le \log\frac{1-\Phi((u(x))}{1-\Phi(x)} &= (x-u(x))\frac{\varphi(\bar{u})}{1-\Phi(\bar{u})}\\ &\le (x-u(x))(Ax+B) \le C\frac{x^3+x^2y+x|y|+x^2}{\sqrt{n}} \le \bar{C}\frac{x^3+x^2|y|+|y|+1}{\sqrt{n}}\end{aligned}$$

with some $u(x) \le \bar{u} \le x$ and appropriate constants $A > 0$, $B > 0$, $C > 0$ and $\bar{C} > 0$. A similar estimation holds for $\log \frac{1-\Phi((v(x))}{1-\Phi(x))}$. These estimates imply that

$$[1-\Phi(x)] \exp\left\{\frac{-C(x^3+x^2|y|+|y|+1}{\sqrt{n}}\right\} \le 1-\Phi(v(x))$$
$$\le 1-\Phi((u(x)) \le [1-\Phi(x)] \exp\left\{\frac{C(x^3+x^2|y|+|y|+1}{\sqrt{n}}\right\}$$

with some appropriate constant $C > 0$. The last estimate together with relation (A4′) imply formula (14).

5b) We can write by formula (13a) and estimates (12c) and (12b)

$$f_n(\varepsilon\sqrt{n}+u|y) \le K \frac{f_n\left(\frac{\varepsilon\sqrt{n}+u+|y|}{\sqrt{2}}\right)}{f_{2n}(y)} \le K \frac{f_n\left(\frac{\varepsilon\sqrt{n}}{\sqrt{2}}\right)}{f_{2n}(y)} e^{-t\sqrt{n}(u+|y|)} \tag{A5}$$

with some appropriate constant $K > 0$ for all $u \ge 0$, where t is the solution of the $\frac{R'(s)}{R(s)} = 2^{-1/2}\varepsilon$, with $R(s) = \int e^{sx} F(\,dx)$. Let us also assume that $|y| \le \frac{\varepsilon}{4}$. In this case the density functions in formula (A5) can be well bounded by means of the result formulated under the name *Sharp form of the local central limit theorem.* For the solution of the present problem it is enough to have the fairly weak estimate $f_n(\varepsilon\sqrt{n}+u|y) \le Ke^{-\varepsilon^2 n/8}e^{-\sqrt{n}u}$ if $\varepsilon > 0$ is chosen sufficiently small with an appropriate constant $K > 0$, and the constant $t = t(\varepsilon)$ in formula (A5) is strictly positive. This implies that

$$0 \le 1 - F_n(\varepsilon\sqrt{n}|y) = \int_0^\infty f_n(\varepsilon\sqrt{n}+u|y)\,dy \le Ke^{-\varepsilon^2 n/8} \le [1-\Phi(x)]e^{-\varepsilon^2 n/20}.$$

This estimate together with formula (14) imply the first relation in formula (10) if $0 \le x \le \frac{\varepsilon}{4}\sqrt{n}$, $|y| \le \frac{\varepsilon}{4}\sqrt{n}$, since it means that the quantity $1 - F_n(\varepsilon\sqrt{n}|y)$ is negligibly small in comparison with $F_n(\varepsilon\sqrt{n}|y) - F_n(x|y)$. The second relation in (10) can be proved similarly, or it follows from the first one if we apply it for the appropriate conditional distributions of the partial sums of the independent and identically distributed random variables $-X_1, -X_2, \dots,$ instead of the partial sum of the original sequence $X_1, X_2, \dots,$ with distribution F.

Appendix. The proof of the sharp form of the local central limit theorem. Let us introduce the functions $\varphi(t) = Ee^{itX_1}$, the characteristic function, and $R(s) = \varphi(-is) = Ee^{sX_1}$, the moment generating function of the random variable X_1. Then the characteristic function of $S_n = \frac{1}{\sqrt{n}} \sum\limits_{k=1}^{n} X_k$ equals $\varphi\left(\frac{t}{\sqrt{n}}\right)^n$, and its moment generating function equals $R\left(\frac{s}{\sqrt{n}}\right)^n$. Put $\psi(s) = \log R(s)$, and let $\psi(z)$ denote its analytic continuation to the plane of the complex numbers in a small neighbourhood of the origin. Such an extension really exists. By the conditions of the result the characteristic function $\varphi\left(\frac{t}{\sqrt{n}}\right)^n$ is integrable for $n \ge n_0$. Hence the inverse Fourier transform can be applied, and it expresses (the existing) density function $f_n(x)$ of S_n. Moreover, the integral expressing the inverse Fourier transform can be replaced to the line $z = t - is\sqrt{n}$ on the plane of complex numbers with a (small) fixed number s. In such a way we can express the density function $f_n(x)$ of S_n by the formula

$$f_n(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \varphi\left(\frac{t}{\sqrt{n}}\right)^n \, dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx - s\sqrt{n}x} \varphi\left(\frac{t}{\sqrt{n}} - is\right)^n \, dt \tag{B1}$$

for all $|s| \le \varepsilon$ if $n \ge k$ with some sufficiently small $\varepsilon > 0$. (Let us remark that because of the existence of the density function $\lim\limits_{|t|\to\infty} \varphi(t) = 0$ by Riemann's lemma. Moreover, the relation $\lim\limits_{|t|\to\infty} \varphi(t+is) = 0$ also holds if $|s| \le \varepsilon$, and the convergence is uniform in the variable s. This observation helps us to justify the above replacement of the integral.)

We want to show that if the parameter s in the integral at the right-hand side of (B1) is appropriately chosen (by means of the saddle point method), then this integral is essentially concentrated in a small neighbourhood of the origin, and this enables us to give a good estimate of the function on the value of $f_n(x)$.

To show this let us observe that $\left|\varphi\left(\frac{t}{\sqrt{n}} - is\right)\right| \le R(s)$. Moreover, for all $\eta > 0$ there exists some $\delta = \delta(\eta) > 0$ such that $\left|\varphi\left(\frac{t}{\sqrt{n}} - is\right)\right| \le (1-\delta)R(s)$ if $|t| \ge \eta\sqrt{n}$. To see the last relation observe that a sum of k independent F distributed random variables $X_1 + \cdots + X_k$, has an integrable characteristic function $\varphi^k(t)$, hence it also has a density function $\bar{f}_k(x)$. This implies that $|\varphi^k(t+is)| < R^k(s)$ with a strict inequality if $|s| \le \varepsilon$ and $t \ne 0$, and also the relation $\lim\limits_{|t|\to\infty} \varphi(t+is) = 0$ holds, moreover it holds uniformly in the parameter s for $|s| \le \varepsilon$. This implies the above formulated inequality. By applying this inequality together with the integrability of the function $\varphi^k(s+it)$ with a fixed s in the variable t, (this statement is equivalent with relation (11)) we get for all $n \ge k$ that

$$\begin{aligned} &\left| \int_{|t| \le \eta\sqrt{n}} e^{-itx - s\sqrt{n}x} \varphi^n\left(\frac{t}{\sqrt{n}} - is\right) dt \right| \\ &\qquad \le (1-\delta)^{n-k} e^{-s\sqrt{n}x} R^{n-k}(s) \int \left| \varphi^k\left(\frac{t}{\sqrt{n}} - is\right) \right| dt \\ &\qquad \le \bar{K}\sqrt{n}(1-\delta)^{n-k} e^{-s\sqrt{n}x} R^{n-k}(s) \le K(1-\delta)^n e^{-s\sqrt{n}x} R^n(s) \end{aligned} \tag{B2}$$

if $|s|\le\varepsilon$ with an appropriate $\delta=\delta(\eta)>0$ and $K=K(\eta)>0$ for all $\eta>0$.

The integrand in the integral at the right-hand side of (B1) can be rewritten as

$$\exp\left\{-\sqrt{n}x\left(s+i\frac{t}{\sqrt{n}}\right)+n\psi\left(s+i\frac{t}{\sqrt{n}}\right)\right\}. \tag{B3}$$

Let us consider the function $s=h(x)$, defined by the equation $x=\psi'(s)$, i.e. $h(\cdot)$ is the inverse function of the function $\psi'(s)=\frac{R'(s)}{R(s)}$. Since $\psi(0)=0$, $\psi'(0)=EX_1^2=1$, $h(x)$ is an analytic function in a small neighbourhood of the origin, $h(0)=0$ and $h'(0)=1$. Let us choose $s=h\left(\frac{x}{\sqrt{n}}\right)$, which function is defined for $|x|\le\varepsilon\sqrt{n}$ with a sufficiently small $\varepsilon>0$. (The saddle point method suggests such a choice of the number s. To explain why the saddle point method suggests such a choice it is useful to introduce a new variable $\bar s=s\sqrt{n}$ and to consider the expression in the exponent of formula (B3) as an analytical function of the argument $z=\bar s+it$. The saddle point method suggests to replace the integral of the analytic function we want to estimate to a new line which goes through a saddle point, i.e. through a point where the derivative of the integrand equals zero in the 'right direction'. If we are looking a saddle point of the special form $z=\bar s$ of the function in the exponent of (B3), then this leads to the equation $\frac{x}{\sqrt{n}}=\psi'\left(\frac{\bar s}{\sqrt{n}}\right)=\psi'(s)$.) We want to give a good estimate on the integral of the function in formula (B3) if we are integrating it in the interval $-\eta\sqrt{n}\le t\le\eta\sqrt{n}$ with a small number $\eta>0$ and a fixed number $s=h\left(\frac{x}{\sqrt{n}}\right)$. To do this let us first consider the Taylor expansion of the function in (B3) as a function of the variable t with a fixed number s. Since the derivative of this function in the origin is zero in the point $t=0$, its second derivative is $-\psi''(s)=-\psi''\left(h\left(\frac{x}{\sqrt{n}}\right)\right)$, (a number close to -1), and its third derivative is of order $O\left(\frac{1}{\sqrt{n}}\right)$ with an order uniform for $|s|\le\varepsilon$, $|t|\le\eta\sqrt{n}$, we get that

$$\begin{aligned}&\exp\left\{-\sqrt{n}x\left(s+i\frac{t}{\sqrt{n}}\right)+n\psi\left(s+i\frac{t}{\sqrt{n}}\right)\right\}\\&\qquad=e^{-\sqrt{n}xs+n\psi(s)}e^{-\psi''(s)t^2/2}\left(1+O\left(\frac{t^3}{\sqrt{n}}\right)\right)\end{aligned}$$

and

$$\begin{aligned}&\int_{-\eta\sqrt{n}}^{\eta\sqrt{n}}e^{-itx-s\sqrt{n}x}\varphi\left(\frac{t}{\sqrt{n}}-is\right)^n\,dt\\&\qquad=e^{-\sqrt{n}xs+n\psi(s)}\int_{-\eta\sqrt{n}}^{\eta\sqrt{n}}e^{-\psi''(s)t^2/2}\left(1+O\left(\frac{t^3}{\sqrt{n}}\right)\right)\,dt\\&\qquad=e^{-\sqrt{n}xs+n\psi(s)}\sqrt{\frac{2\pi}{\psi''(s)}}\left(1+O\left(\frac{1}{\sqrt{n}}\right)\right).\end{aligned} \tag{B4}$$

Relations (B1), (B2) and (B4) together with the relation $R^n(s)=e^{n\psi(s)}$ imply that

$$f_n(x)=\frac{e^{-n(sx/\sqrt{n}-\psi(s))}}{\sqrt{2\pi\psi''(s)}}\left(1+O\left(\frac{1}{\sqrt{n}}\right)\right) \tag{B5}$$

with $s=h\left(\frac{x}{\sqrt{n}}\right)$ if $|x|\le\varepsilon\sqrt{n}$ with a sufficiently small $\varepsilon>0$.

Both $H_1\left(\frac{x}{\sqrt{n}}\right)=\frac{sx}{\sqrt{n}}-\psi(s))$ and $H_2\left(\frac{x}{\sqrt{n}}\right)=\psi''(s)$ are analytic functions of the variable $\frac{x}{\sqrt{n}}$ after the substitution $s=h\left(\frac{x}{\sqrt{n}}\right)$ which do not depend on the parameter n. Beside this, we can write $s=\frac{x}{\sqrt{n}}+\frac{x}{\sqrt{n}}A\left(\frac{x}{\sqrt{n}}\right)$, and $\psi(s)=\frac{s^2}{2}+s^3B(s)$ with some analytic functions $A(\cdot)$ and $B(\cdot)$. Hence $-n(sx/\sqrt{n}-\psi(s))=-\frac{x^2}{2}+\frac{x^3}{\sqrt{n}}\lambda\left(\frac{x}{\sqrt{n}}\right)$ with some appropriate analytic function $\lambda(\cdot)$ in a small neighborhood of the origin. Similarly, $\psi''(s)=1+sC(s)$ with an analytic function $C(s)$ in a small neighbourhood of the origin, hence we can write $\psi''(s)=1+\frac{x}{\sqrt{n}}\mu\left(\frac{x}{\sqrt{n}}\right)$ with an appropriate analytical function $\mu(x)$. The first relation of formula (12a) follows from formula (B5) and the observation made about the expressions in this formula made after it. The second relation of formula (12a) is a simple consequence of the first one.

We could prove the asymptotic relation (12a) only for such numbers x for which the equation $\frac{x}{\sqrt{n}}=\psi'(s)$ has a solution, and we could guarantee it only for $|x|\le\varepsilon\sqrt{n}$ with some $\varepsilon>0$. In the general case we can give a sharp asymptotic formula of the density function $f_n(x)$ only for such arguments x. To get a good upper bound for the density function $f_n(\bar x)$ in the case of a general number $\bar x$ let us write this number in the form $\bar x=x+z$ with some $|x|\le\varepsilon\sqrt{n}$, and let us express $f_n(\bar x)$ by formula (B1) (with the replacement of x by $\bar x$ in it) with the same number s, as before, i.e. let us choose the parameter s as the solution of the equation $\frac{x}{\sqrt{n}}=\psi'(s)$. The estimate (B2) remains valid if we replace x by $\bar x$ everywhere in this formula. To get an appropriate upper bound in formula (B4) in the new situation let us first give a good upper bound on the expression in formula (B3). Let us recall that $|e^z|=e^{\operatorname{Re} z}$.

We can get, by means of a Taylor expansion that

$$\operatorname{Re}\left(-\sqrt{n}\bar x\left(s+i\frac{t}{\sqrt{n}}\right)+n\psi\left(s+i\frac{t}{\sqrt{n}}\right)\right)=-\sqrt{n}s\bar x+n\psi(s)-\psi''(s)\frac{t^2}{2}+O\left(\frac{t^3}{\sqrt{n}}\right)$$

if $|t|\le\eta\sqrt{n}$ with some appropriately small $\eta>0$. Hence we get the following analog of relation (B4):

$$\left|\int_{-\eta\sqrt{n}}^{\eta\sqrt{n}}e^{-it\bar x-s\sqrt{n}\bar x}\varphi\left(\frac{t}{\sqrt{n}}-is\right)^n\,dt\right|$$
$$\le e^{-\sqrt{n}\bar x s+n\psi(s)}\sqrt{\frac{2\pi}{\psi''(s)}}\left(1+O\left(\frac{1}{\sqrt{n}}\right)\right). \tag{B4$'$}$$

Relations (B4$'$) together with the version of relations (B1) and (B3) (for $\bar x$ instead of x imply that

$$f_n(\bar x)\le \bar K e^{-\sqrt n\bar x s+n\psi(s)}\frac{1}{\sqrt{2\pi\psi''(s)}}$$

with an appropriate constant $\bar K>0$. A comparison of this formula with relation (B5) yields that

$$f_n(\bar x)\le Ke^{-\sqrt n s(\bar x-x)}\frac{e^{-\sqrt n xs+n\psi(s)}}{\sqrt{2\pi\psi''(s)}}\le Ke^{-\sqrt n s(\bar x-x)}f_n(x)=Ke^{-\sqrt n sz}f_n(x)$$

with some appropriate constant $K>0$, as we claimed in formula (12b).

Finally relation (12c) is a simple consequence of relation (12b) with the choice $x=0$ if we observe that $s=0$ in this case, and $f_n(0)$ is bounded by a constant not depending on n by formula (12a). Actually, a sharper estimate holds. With some extra-work it can be shown that $f_n(x)$ can be approximated by the standard normal density function $\varphi(x)$ with an error bounded by $\frac{\text{const.}}{\sqrt n}$ in the supremum norm. But we do not need such an estimate, hence its proof will be omitted.