# Elementary limit theorems in probability

Jason Swanson

December 27, 2008

## 1 Introduction

What follows is a collection of various limit theorems that occur in probability. Most are taken from a short list of references. Such theorems are stated without proof and a citation follows the name of the theorem. A few are not taken from references. They are usually straightforward generalizations of the standard theorems and proofs are provided. The glossary includes some definitions of terms used in these theorems.

## 2 Glossary

**uncorrelated:** A family of random variables $\{X_i\}_{i \in I}$ with $EX_i^2 < \infty$ is uncorrelated if $E(X_i X_j) = EX_i EX_j$ whenever $i \neq j$.

**independent:** The random variables $X_1, X_2, \ldots, X_n$ are independent if

$$P\left(\bigcap_{j=1}^n \{X_j \in A_j\}\right) = \prod_{j=1}^n P(X_j \in A_j)$$

for all $n$-tuples of measurable sets $(A_1, A_2, \ldots, A_n)$. A family of random variables $\{X_i\}_{i \in I}$ is independent if for each finite subset $J \subset I$, the family $\{X_i\}_{i \in J}$ is independent.

**uniformly integrable:** A family of random variables $\{X_i\}_{i \in I}$ is uniformly integrable if $\sup_{i \in I} E\left[|X_i| 1_{\{|X_i| \geq K\}}\right] \to 0$ as $K \to \infty$.

**pairwise independent:** A family of random variables $\{X_i\}_{i \in I}$ is pairwise independent if $X_i$ and $X_j$ are independent whenever $i \neq j$.

**measure preserving:** Let $(\Omega, \mathcal{F}, P)$ be a probability space. An injective map $T : \Omega \to \Omega$ is a measure preserving transformation if $A \in \mathcal{F}$ implies $T(A) \in \mathcal{F}$ and $P(T(A)) = P(A)$.

**$T$-invariant:** If $T$ is a measure preserving transformation, then a set $\Lambda \in \mathcal{F}$ is $T$-invariant, or invariant under $T$, if $1_\Lambda(\omega) = 1_\Lambda(T(\omega))$ a.s.

**median:** Let $x_1, \ldots, x_n \in \mathbb{R}$ and let $\tau : \{1, \ldots, n\} \to \{1, \ldots, n\}$ be a bijection such that $y_j = x_{\tau(j)}$ satisfies $y_1 \leq \cdots \leq y_n$. Then $\operatorname{med}(x_1, \ldots, x_n) = y_k$, where $k = \lfloor (n+1)/2 \rfloor$.

If $x^{(1)}, \ldots, x^{(n)} \in \mathbb{R}^d$, then $\text{med}(x^{(1)}, \ldots, x^{(n)})$ is the vector in $\mathbb{R}^d$ whose $j$-th component is $\text{med}(x_j^{(1)}, \ldots, x_j^{(n)})$.

**weak convergence:** Let $\mu_n$, $\mu$ be probability measures on $(S, \mathcal{B})$, where $S$ is a metric space and $\mathcal{B}$ is its Borel $\sigma$-algebra. If $\int f\, d\mu_n \to \int f\, d\mu$ for every bounded, continuous $f : S \to \mathbb{R}$, then $\mu_n$ converges weakly to $\mu$ (written $\mu_n \Rightarrow \mu$).

**convergence in distribution:** Let $X_n$, $X$ be random variables taking values in a metric space $S$. Define measures $\mu_n$, $\mu$ on $(S, \mathcal{B})$, where $\mathcal{B}$ is the Borel $\sigma$-algebra on $S$, by $\mu_n(A) = P(X_n \in A)$, $\mu(A) = P(X \in A)$. Then $X_n$ converges to $X$ in distribution (written $X_n \xrightarrow{d} X$ or $X_n \Rightarrow X$) if $\mu_n \Rightarrow \mu$.

**multinormal:** A random vector $X = (X_1, \ldots, X_d)^T$ is multinormal if every linear combination $c_1X_1 + \cdots + c_dX_d$ has a normal (possibly degenerate) distribution. The mean of $X$ is the (column) vector $\theta \in \mathbb{R}^d$ with $\theta_j = EX_j$. The covariance of $X$ is the $d \times d$ matrix $\sigma$ with

$$\sigma_{ij} = \text{Cov}(X_i, X_j) = E(X_iX_j) - \theta_i\theta_j.$$

In this case, $\sigma$ is symmetric and positive semidefinite, and $c_1X_1 + \ldots + c_dX_d$ has mean $c^T\theta$ and variance $c^T\sigma c$.

**Poisson:** A random variable $Z$ is Poisson($\lambda$) if $P(Z = k) = e^{-\lambda}\lambda^k/k!$ for all $k \in \mathbb{Z}_+ = \{0\} \cup \mathbb{N}$.

**stable law:** A random variable $Y$ has a stable law if for every $k \in \mathbb{N}$, there are constants $a_k$ and $b_k$ such that $(Y_1 + \cdots + Y_k - b_k)/a_k \stackrel{d}{=} Y$ whenever $Y_1, \ldots, Y_k$ are independent and identically distributed (iid) with $Y_j \stackrel{d}{=} Y$. (The notation $U \stackrel{d}{=} V$ means that the random variables $U$ and $V$ have the same distribution.)

**slowly varying:** A function $L(x)$ is slowly varying if $\lim_{x\to\infty} L(tx)/L(x) = 1$ for all $t > 0$.

# 3 Laws of Large Numbers

**Theorem 3.1.** ($L^2$ weak law)[2]
*Let $X_1, X_2, \ldots$ be uncorrelated random variables with $EX_i = \mu$ and $\text{Var}(X_i) \le C < \infty$. If $S_n = X_1 + \cdots + X_n$, then $S_n/n \to \mu$ in $L^2$ and in probability as $n \to \infty$.*

**Theorem 3.2.** ($L^1$ weak law)[7]
*If $X_1, X_2, \ldots$ is a uniformly integrable sequence of independent random variables, then*

$$\frac{1}{n}\sum_{m=1}^{n}(X_m - EX_m) \to 0,$$

*in $L^1$ and in probability as $n \to \infty$. In particular, if $X_1, X_2, \ldots$ is iid with $E|X_1| < \infty$ and $S_n = X_1 + \cdots + X_n$, then $S_n/n \to EX_1$ in $L^1$ and in probability.*

**Theorem 3.3.** (Weak law for triangular arrays)[2]
*For each $n$, let $X_{n,k}$, $1 \le k \le n$, be independent. Let $b_n > 0$ with $b_n \to \infty$ and let $\overline{X}_{n,k} = X_{n,k} 1_{\{|X_{n,k}| \le b_n\}}$. Suppose that*

*(i)* $\sum_{k=1}^n P(|X_{n,k}| > b_n) \to 0$, *and*

*(ii)* $b_n^{-2} \sum_{k=1}^n E\overline{X}_{n,k}^2 \to 0$ *as* $n \to \infty$.

*If we let $S_n = X_{n,1} + \cdots + X_{n,n}$ and put $a_n = \sum_{k=1}^n E\overline{X}_{n,k}$, then $(S_n - a_n)/b_n \to 0$ in probability.*

**Theorem 3.4.** (Weak law of large numbers)
*Let $X_1, X_2, \ldots$ be independent with $\sup_k xP(|X_k| \ge x) \to 0$ as $x \to \infty$. Let $S_n = X_1 + \cdots + X_n$ and let*

$$\mu_n = \frac{1}{n} \sum_{k=1}^n E[X_k 1_{\{|X_k| \le n\}}].$$

*Then $S_n/n - \mu_n \to 0$ in probability.*

**Proof.** Let $f(x) = \sup_k xP(|X_k| \ge x)$. We apply Theorem 3.3 with $X_{n,k} = X_k$ and $b_n = n$. To verify Condition (i) of Theorem 3.3, note that

$$\sum_{k=1}^n P(|X_k| > n) \le n \sup_k P(|X_k| > n) = f(n) \to 0,$$

as $n \to \infty$. For Condition (ii), we use the fact that for any random variable $Y$, we have $EY^2 = \int_0^\infty 2yP(|Y| > y)\, dy$. Thus,

$$\frac{1}{n^2} \sum_{k=1}^n E\overline{X}_{n,k}^2 \le \frac{2}{n^2} \sum_{k=1}^n \int_0^n yP(|X_k| > y)\, dy \le \frac{2}{n} \int_0^n f(y)\, dy.$$

Fix $\varepsilon > 0$. Since $f(x) \to 0$ as $x \to \infty$, there exists $K > 0$ such that $x \ge K$ implies $f(x) \le \varepsilon$. Thus,

$$\begin{aligned} \limsup_{n\to\infty} \frac{2}{n} \int_0^n f(y)\, dy &= \limsup_{n\to\infty} \left( \frac{2}{n} \int_0^K f(y)\, dy + \frac{2}{n} \int_K^n f(y)\, dy \right) \\ &\le \limsup_{n\to\infty} \left( \frac{2K^2}{n} + \frac{2(n-K)\varepsilon}{n} \right) = 2\varepsilon. \end{aligned}$$

Since $\varepsilon$ was arbitrary, this completes the proof. □

**Theorem 3.5.** (Strong laws of large numbers)[2],[4]

*(i) If $X_1, X_2, \ldots$ are pairwise independent and identically distributed with $E|X_1| < \infty$, then $(X_1 + \cdots + X_n)/n \to EX_1$ a.s. as $n \to \infty$.*

*(ii) If $X_1, X_2, \ldots$ are iid with $E[|X_1| 1_{\{X_1 > 0\}}] = \infty$ and $E[|X_1| 1_{\{X_1 < 0\}}] < \infty$, then $(X_1 + \cdots + X_n)/n \to \infty$ a.s.*

*(iii) If $X_1, X_2, \dots$ are iid with $EX_1^2 < \infty$, then $(X_1 + \cdots + X_n)/n \to EX_1$ a.s. and in $L^2$.*

Part (i) of the above theorem is Theorem 1.7.1 in [2]. If we inspect the proof of that theorem, we see that the conditions can be weakened to the following.

**Theorem 3.6.** (Generalized strong law of large numbers)
*Let $X_1, X_2, \dots$ be nonnegative random variables with $E|X_k| < \infty$ and $EX_k = \mu$ for all $k$. Let $Y_k = X_k 1_{\{|X_k| \le k\}}$, and assume the random variables $\{Y_k\}$ are uncorrelated. Also assume there exists a constant $C$ such that $P(X_k > t) \le CP(X_1 > t)$ for all $k$ and $t$. If $S_n = X_1 + \cdots + X_n$, then $S_n/n \to \mu$ a.s. as $n \to \infty$.*

*Remark* 3.7. Any sequence whose positive and negative parts satisfy the above hypotheses will satisfy the strong law of large numbers.

**Proof.** Let $\alpha > 1$ and $\varepsilon > 0$ be arbitrary. Let $T_n = Y_1 + \cdots + Y_n$ and $k(n) = \lfloor \alpha^n \rfloor$. By Chebyshev,

$$\begin{aligned}\sum_{n=1}^{\infty} P(|T_{k(n)} - ET_{k(n)}| > \varepsilon k(n)) &\le \varepsilon^{-2} \sum_{n=1}^{\infty} \mathrm{Var}(T_{k(n)})/k(n)^2 \\ &= \varepsilon^{-2} \sum_{n=1}^{\infty} k(n)^{-2} \sum_{m=1}^{k(n)} \mathrm{Var}(Y_m) \\ &= \varepsilon^{-2} \sum_{m=1}^{\infty} \mathrm{Var}(Y_m) \sum_{n: k(n) \ge m} k(n)^{-2}.\end{aligned}$$

Since $\lfloor \alpha^n \rfloor \ge \alpha^n/2$,

$$\sum_{n: \lfloor \alpha^n \rfloor \ge m} \lfloor \alpha^n \rfloor^{-2} \le 4 \sum_{n: \lfloor \alpha^n \rfloor \ge m} \alpha^{-2n} \le 4(1 - \alpha^{-2}) m^{-2}.$$

Hence,

$$\sum_{n=1}^{\infty} P(|T_{k(n)} - ET_{k(n)}| > \varepsilon k(n)) \le 4(1 - \alpha^{-2}) \varepsilon^{-2} \sum_{m=1}^{\infty} \mathrm{Var}(Y_m)/m^2.$$

To bound this sum, note that

$$\mathrm{Var}(Y_k) \le EY_k^2 = \int_0^{\infty} 2yP(Y_k > y)\, dy \le \int_0^k 2yP(X_k > y)\, dy \le C \int_0^k 2yP(X_1 > y)\, dy.$$

Thus,

$$\begin{aligned}\sum_{m=1}^{\infty} \mathrm{Var}(Y_m)/m^2 &\le C \sum_{m=1}^{\infty} m^{-2} \int_0^{\infty} 1_{\{y < m\}} 2yP(X_1 > y)\, dy \\ &= C \int_0^{\infty} \left\{ \sum_{m=1}^{\infty} m^{-2} 1_{\{y < m\}} \right\} 2yP(X_1 > y)\, dy \\ &\le 4C \int_0^{\infty} P(X_1 > y)\, dy = 4CE|X_1|,\end{aligned}$$

where the last inequality above uses Lemma 1.7.1(c) in [2]. Putting this together, we have

$$\sum_{n=1}^{\infty} P(|T_{k(n)} - ET_{k(n)}| > \varepsilon k(n)) \leq 16C(1-\alpha^{-2})\varepsilon^{-2}E|X_1| < \infty.$$

By the Borel-Cantelli lemma, since $\varepsilon$ is arbitrary, this implies $(T_{k(n)} - ET_{k(n)})/k(n) \to 0$ a.s. Now,

$$EY_k = E(X_k - X_k 1_{\{X_k > k\}}) = \mu - \int_0^\infty P(X_k 1_{\{X_k>k\}} > y)\, dy.$$

Note that

$$\begin{aligned}\int_0^\infty P(X_k 1_{\{X_k>k\}} > y)\, dy &= \int_0^k P(X_k > k)\, dy + \int_k^\infty P(X_k > y)\, dy \\ &\leq C\left(\int_0^k P(X_1 > k)\, dy + \int_k^\infty P(X_1 > y)\, dy\right) \\ &= CE(X_1 1_{\{X_1 > k\}}) \to 0.\end{aligned}$$

Hence, $EY_k \to \mu$ as $k \to \infty$, which implies $ET_{k(n)}/k(n) \to \mu$. We have therefore shown that $T_{k(n)}/k(n) \to \mu$ a.s. For the intermediate values, if $k(n) \leq m < k(n+1)$, then

$$\frac{T_{k(n)}}{k(n+1)} \leq \frac{T_m}{m} \leq \frac{T_{k(n+1)}}{k(n)},$$

where we have used the fact that $Y_k \geq 0$. Thus, recalling that $k(n) = \lfloor \alpha^n \rfloor$, we have $k(n+1)/k(n) \to \alpha$ and

$$\frac{1}{\alpha}\mu \leq \liminf_{n\to\infty} T_m/m \leq \limsup_{n\to\infty} T_m/m \leq \alpha\mu.$$

Since $\alpha > 1$ was arbitrary, this shows that $T_m/m \to \mu$ a.s.

Finally, note that

$$\sum_{k=1}^{\infty} P(X_k > k) \leq C\sum_{k=1}^{\infty} P(X_1 > k) \leq C\int_0^\infty P(X_1 > y)\, dy = CEX_1 < \infty.$$

By Borel-Cantelli, $P(X_k \neq Y_k \text{ i.o.}) = 0$. Therefore, $|S_n(\omega) - T_n(\omega)| \leq R(\omega) < \infty$ a.s. for all $n$, which implies $S_n/n \to \mu$ a.s. □

**Theorem 3.8.** (Kolmogorov's strong law of large numbers)[4]
*Let $X_1, X_2, \ldots$ be iid and let $S_n = \sum_{j=1}^n X_j$. Then there exists $\mu \in \mathbb{R}$ such that $\lim_{n\to\infty} S_n/n = \mu$ a.s. if and only if $E|X_1| < \infty$. In this case, $\mu = EX_1$.*

**Theorem 3.9.** (Glivenko-Cantelli theorem)[2]
*Suppose $X_1, X_2, \ldots$ are iid. Define $F(x) = P(X_1 \leq x)$ and $F_n(x) = n^{-1}\sum_{j=1}^n 1_{\{X_j \leq x\}}$. Then $\sup_{x\in\mathbb{R}} |F_n(x) - F(x)| \to 0$ a.s.*

**Theorem 3.10.** (Ergodic strong law of large numbers)[4]
*Let $T$ be an injective measure preserving transformation of $\Omega$ onto itself. Assume the only $T$-invariant sets are sets of probability 0 or 1. If $X \in L^1$, then $n^{-1}\sum_{j=1}^n X(T^j(\omega)) \to EX$ a.s. and in $L^1$ as $n \to \infty$, where $T^{j+1} = T^j \circ T$.*

**Theorem 3.11.** (Quantile strong law of large numbers)
*Let $X_1, X_2, \ldots$ be iid and let $F(x) = P(X_1 \le x)$. Fix $\alpha \in (0,1)$. Suppose there exists $q$ such that $F(q) = \alpha$ and $F$ is strictly increasing at $x = q$. Let $M_n$ denote the $\lfloor \alpha n \rfloor$-th order statistic of the quantities $X_1, \ldots, X_n$. Then $M_n \to q$ a.s.*

**Proof.** Note that for each $x \in \mathbb{R}$, $\{M_n \le x\} = \{\sum_{j=1}^n 1_{\{X_j \le x\}} \ge \lfloor \alpha n \rfloor\}$. Thus, for all $\varepsilon > 0$,

$$\{q - \varepsilon < M_n \le q + \varepsilon\} = \left\{\frac{1}{n}\sum_{j=1}^n 1_{\{X_j \le q+\varepsilon\}} \ge \frac{\lfloor \alpha n \rfloor}{n}\right\} \cap \left\{\frac{1}{n}\sum_{j=1}^n 1_{\{X_j \le q-\varepsilon\}} < \frac{\lfloor \alpha n \rfloor}{n}\right\}.$$

By Theorem 3.9, there exists $\Omega^* \subset \Omega$ such that $P(\Omega^*) = 1$ and $n^{-1}\sum_{j=1}^n 1_{\{X_j \le x\}} \to F(x)$ uniformly in $x$ for each $\omega \in \Omega^*$. Since $F(q+\varepsilon) > \alpha$ and $F(q-\varepsilon) < \alpha$, each $\omega \in \Omega^*$ is an element of $\{q - \varepsilon < M_n \le q + \varepsilon\}$ for sufficiently large $n$. Hence,

$$\Omega^* \subset \{q - \varepsilon \le \liminf_{n\to\infty} M_n \le \limsup_{n\to\infty} M_n \le q + \varepsilon\},$$

which shows that $M_n \to q$ a.s. □

*Remark* 3.12. There are many very detailed results on the asymptotics of order statistics in [5].

# 4 Convergence in Distribution

**Theorem 4.1.** (Portmanteau Theorem) [1]
*If $X_n$, $X$ are random variables taking values in a metric space $S$, then the following are equivalent:*

*(i)* $X_n \Rightarrow X$

*(ii)* $E[f(X_n)] \to E[f(X)]$ *for all bounded, uniformly continuous* $f : S \to \mathbb{R}$

*(iii)* $\limsup_{n\to\infty} P(X_n \in F) \le P(X \in F)$ *for all closed* $F \subset S$

*(iv)* $\liminf_{n\to\infty} P(X_n \in G) \le P(X \in G)$ *for all open* $G \subset S$

*(v)* $\lim_{n\to\infty} P(X_n \in A) = P(X \in A)$ *for all Borel sets* $A \subset S$ *with* $P(X \in \partial A) = 0$

**Theorem 4.2.** (Skorohod representation)[3]
*Let $S$ be a complete, separable metric space and $\mathcal{B}$ its Borel $\sigma$-algebra. If $\mu_n$, $\mu_0$ are probability measures on $(S, \mathcal{B})$ with $\mu_n \Rightarrow \mu_0$, then there exists a probability space $(\Omega, \mathcal{F}, P)$ and random variables $X_n$, $X_0$ on $\Omega$ taking values in $S$ such that $X_n$ has distribution $\mu_n$ for all $n \ge 0$ and $X_n \to X_0$ a.s.*

**Theorem 4.3.** [1]
*Let $S$ be a metric space, $\mathcal{B}$ its Borel $\sigma$-algebra, and $P_n$, $P$ probability measures on $(S, \mathcal{B})$. Suppose $\mathcal{U} \subset \mathcal{B}$ satisfies*

*(i) $\mathcal{U}$ is closed under finite intersections, and*

*(ii) each open set in $S$ is a countable union of elements of $\mathcal{U}$.*

*If $P_n(A) \to P(A)$ for all $A \in \mathcal{U}$, then $P_n \Rightarrow P$.*

**Theorem 4.4.** [1]
*For $x, y \in \mathbb{R}^d$, write $x \leq y$ if $x_j \leq y_j$ for all $j$. If $X^{(n)}$, $X$ are $\mathbb{R}^d$-valued random variables with $F^{(n)}(x) = P(X^{(n)} \leq x)$ and $F(x) = P(X \leq x)$, then $X^{(n)} \Rightarrow X$ if and only if $F^{(n)}(x) \to F(x)$ for all $x \in \mathbb{R}^d$ such that $F$ is continuous at $x$.*

**Theorem 4.5.** (Cramér-Wold device)[2]
*Let $X^{(n)}$, $X$ be random vectors in $\mathbb{R}^d$. If $\theta \cdot X^{(n)} \Rightarrow \theta \cdot X$ for all $\theta \in \mathbb{R}^d$, then $X^{(n)} \Rightarrow X$.*

# 5 Central Limit Theorems

**Theorem 5.1.** (Central Limit Theorem)[2]
*Let $X_1, X_2, \ldots$ be iid with $EX_j = \mu$. Suppose that $\mathrm{Var}(X_j) = \sigma^2 \in (0, \infty)$. If $S_n = X_1 + \ldots + X_n$, then $(S_n - n\mu)/(\sigma n^{1/2}) \Rightarrow \chi$, where $\chi$ has the standard normal distribution.*

**Theorem 5.2.** (Lindeberg-Feller Theorem)[2]
*For each $n$, let $X_{n,m}$, $1 \leq m \leq n$, be independent random variables with $EX_{n,m} = 0$. If*

*(i) $\sum_{m=1}^n EX_{n,m}^2 \to \sigma^2$ as $n \to \infty$, and*

*(ii) for each $\varepsilon > 0$, $\sum_{m=1}^n E[|X_{n,m}|^2 1_{\{|X_{n,m}| > \varepsilon\}}] \to 0$ as $n \to \infty$,*

*then $S_n = X_{n,1} + \cdots + X_{n,n} \Rightarrow \sigma\chi$ as $n \to \infty$, where $\chi$ has the standard normal distribution.*

*Remark* 5.3. Durrett assumes in condition (i) that $\sigma^2 > 0$. However, if $\sigma^2 = 0$, then condition (i) says that $S_n \to 0$ in $L^2$ and therefore $S_n \to 0$ in probability and in distribution.

**Theorem 5.4.** (Lyapunov's Central Limit Theorem)
*Let $X_1, X_2, \ldots$ be independent with $EX_j = 0$ for all $j$. Let $\alpha_n = \sqrt{\sum_{j=1}^n \mathrm{Var}(X_j)}$. If there exists $\delta > 0$ such that $\lim_{n\to\infty} \alpha_n^{-(2+\delta)} \sum_{j=1}^n E|X_j|^{2+\delta} = 0$, then*

$$\frac{X_1 + \cdots + X_n}{\sqrt{\sum_{j=1}^n \mathrm{Var}(X_j)}} \Rightarrow \chi,$$

*where $\chi$ has the standard normal distribution.*

**Proof.** For each $n \in \mathbb{N}$ and $m \in \{1, \ldots, n\}$, let $X_{n,m} = \alpha_n^{-1} X_m$. Note that $\sum_{m=1}^{n} EX_{n,m}^2 = 1$. Also note that for each $\varepsilon > 0$,

$$\begin{aligned}\sum_{m=1}^{n} E[|X_{n,m}|^2 1_{\{|X_{n,m}|>\varepsilon\}}] &= \alpha_n^{-2} \sum_{m=1}^{n} E[|X_m|^2 1_{\{|X_m|>\alpha_n\varepsilon\}}] \\ &\le \alpha_n^{-2} \sum_{m=1}^{n} E\left[\frac{|X_m|^{2+\delta}}{(\alpha_n\varepsilon)^\delta} 1_{\{|X_m|>\alpha_n\varepsilon\}}\right] \\ &\le \varepsilon^{-\delta}\alpha_n^{-(2+\delta)} \sum_{m=1}^{n} E|X_m|^{2+\delta} \to 0,\end{aligned}$$

as $n \to \infty$. Hence, by Theorem 5.2, $X_{n,1} + \cdots + X_{n,n} \Rightarrow \chi$. □

**Theorem 5.5.** (Nonclassical Central Limit Theorem, Part I)[6]
*For each $n$, let $X_{n,m}$, $1 \le m \le n$, be independent random variables with $EX_{n,m} = 0$ and $\sigma_{nm}^2 = EX_{n,m}^2 > 0$. Suppose $\sum_{m=1}^{n} \sigma_{nm}^2 = 1$ and let $S_n = X_{n,1} + \cdots + X_{n,n}$. Then $S_n \Rightarrow \chi$, where $\chi$ has the standard normal distribution, if and only if*

$$\lim_{n\to\infty} \sum_{m=1}^{n} \int_{\{|x|>\varepsilon\}} |x| \left|P(X_{n,m} \le x) - P(\sigma_{nm}^{-1}\chi \le x)\right| dx = 0, \tag{5.1}$$

*for every $\varepsilon > 0$.*

**Theorem 5.6.** (Nonclassical Central Limit Theorem, Part II)[6]
*For each $n$, let $X_{n,m}$, $1 \le m \le n$, be independent random variables with $EX_{n,m} = 0$ and $\sigma_{nm}^2 = EX_{n,m}^2 > 0$. Suppose $\sum_{m=1}^{n} \sigma_{nm}^2 = 1$. If Condition (ii) of Theorem 5.2 holds, then (5.1) holds for every $\varepsilon > 0$. Conversely, if (5.1) holds for every $\varepsilon > 0$, and $\max_{1\le m\le n} \sigma_{nm}^2 \to 0$ as $n \to \infty$, then Condition (ii) of Theorem 5.2 holds.*

**Theorem 5.7.** (Converse of Lindeberg-Feller Theorem)
*For each $n$, let $X_{n,m}$, $1 \le m \le n$, be independent random variables with $EX_{n,m} = 0$ and $EX_{n,m}^2 > 0$. Suppose*

*(i)* $\sum_{m=1}^{n} EX_{n,m}^2 \to \sigma^2 > 0$ *as* $n \to \infty$, *and*

*(ii)* $\max_{1\le m\le n} EX_{n,m}^2 \to 0$ *as* $n \to \infty$.

*If $S_n = X_{n,1} + \cdots + X_{n,n} \Rightarrow \sigma\chi$ as $n \to \infty$, where $\chi$ has the standard normal distribution, then Condition (ii) of Theorem 5.2 holds.*

**Proof.** Let $\sigma_n^2 = \sum_{m=1}^{n} EX_{n,m}^2$ and define $\widetilde{X}_{n,m} = \sigma_n^{-1} X_{n,m}$ and $\widetilde{S}_n = \widetilde{X}_{n,1} + \cdots + \widetilde{X}_{n,n}$. Since $S_n \Rightarrow \sigma\chi$ and $\sigma_n \to \sigma$, it follows that $\widetilde{S}_n \Rightarrow \chi$. Hence, by Theorem 5.5,

$$\lim_{n\to\infty} \sum_{m=1}^{n} \int_{\{|x|>\varepsilon\}} |x| \left|P(\widetilde{X}_{n,m} \le x) - P(\sigma_{nm}^{-1}\chi \le x)\right| dx = 0,$$

for every $\varepsilon > 0$. Also, if $\widetilde{\sigma}_{nm}^2 = E\widetilde{X}_{n,m}^2$, then $\widetilde{\sigma}_{nm}^2 > 0$, $\sum_{m=1}^n \widetilde{\sigma}_{nm}^2 = 1$, and $\max_{1\le m\le n} \widetilde{\sigma}_{nm}^2 \to 0$ as $n \to \infty$. Hence, by Theorem 5.6,

$$\lim_{n\to\infty} \sum_{m=1}^{n} E[|\widetilde{X}_{n,m}|^2 1_{\{|\widetilde{X}_{n,m}|>\widetilde{\varepsilon}\}}] = 0,$$

for every $\widetilde{\varepsilon} > 0$.

Note that

$$\sum_{m=1}^{n} E[|X_{n,m}|^2 1_{\{|X_{n,m}|>\varepsilon\}}] = \sigma_n^2 \sum_{m=1}^{n} E[|\widetilde{X}_{n,m}|^2 1_{\{|\widetilde{X}_{n,m}|>\sigma_n^{-1}\varepsilon\}}]$$

Since $\sigma_n \to \sigma$, there exists $C > 0$ such that $\sigma_n \le C$ for all $n$. Hence, if we define $\widetilde{\varepsilon} = C^{-1}\varepsilon$, then

$$\sum_{m=1}^{n} E[|X_{n,m}|^2 1_{\{|X_{n,m}|>\varepsilon\}}] \le C^2 \sum_{m=1}^{n} E[|\widetilde{X}_{n,m}|^2 1_{\{|\widetilde{X}_{n,m}|>\widetilde{\varepsilon}\}}],$$

which tends to zero as $n \to \infty$. □

**Theorem 5.8.** (Berry-Esseen)[2]
*Let $X_1, X_2, \ldots$ be iid with $EX_i = 0$ and $EX_i^2 = \sigma^2 < \infty$. Let $S_n = X_1 + \cdots + X_n$. If $F_n(x) = P(S_n/(\sigma n^{1/2}) \le x)$ and $\Phi(x) = P(\chi \le x)$, then*

$$\sup_x |F_n(x) - \Phi(x)| \le \frac{3\rho}{\sigma^3 n^{1/2}},$$

*where $\rho = E|X_i|^3$.*

**Theorem 5.9.** (The Central Limit Theorem in $\mathbb{R}^d$)[2]
*Let $X^{(1)}, X^{(2)}, \ldots$ be iid random vectors in $\mathbb{R}^d$ with $EX^{(n)} = \mu$ and finite covariances*

$$\sigma_{ij} = E[(X_i^{(n)} - \mu_i)(X_j^{(n)} - \mu_j)].$$

*If $S^{(n)} = X^{(1)} + \cdots + X^{(n)}$, then $(S^{(n)} - n\mu)/n^{1/2} \Rightarrow N$, where $N$ is multinormal with mean $0$ and covariance $\sigma$.*

**Theorem 5.10.** (Multidimensional Lindeberg-Feller Theorem)
*For each $n$, let $X^{(n,m)}$, $1 \le m \le n$, be independent, $\mathbb{R}^d$-valued random vectors with $EX^{(n,m)} = 0$. Let $\sigma^{(n,m)} = (\sigma_{ij}^{(n,m)})$, where $\sigma_{ij}^{(n,m)} = EX_i^{(n,m)} X_j^{(n,m)}$. If*

*(i)* $\sum_{m=1}^n \sigma^{(n,m)} \to \sigma$ *as* $n \to \infty$, *and*

*(ii) for each $\theta \in \mathbb{R}^d$ and each $\varepsilon > 0$, $\sum_{m=1}^n E[|\theta \cdot X^{(n,m)}|^2 1_{\{|\theta\cdot X^{(n,m)}|>\varepsilon\}}] \to 0$ as $n \to \infty$,*

*then $S^{(n)} = X^{(n,1)} + \cdots + X^{(n,n)} \Rightarrow N$, where $N$ is multinormal with mean $0$ and covariance $\sigma$.*

**Proof.** Fix $\theta \in \mathbb{R}^d$. By the Cramér-Wold device, it suffices to show $\theta \cdot S^{(n)} \Rightarrow \theta \cdot N$. Now, for each $n, m \in \mathbb{N}$ with $1 \le m \le n$, let $Y_{n,m} = \theta \cdot X^{(n,m)}$. Then

(a) $Y_{n,m}$, $1 \le m \le n$, are independent,

(b) $EY_{n,m} = 0$, and

(c) $\sum_{m=1}^n EY_{n,m}^2 = \sum_{m=1}^n E|\theta \cdot X^{(n,m)}|^2 = \sum_{m=1}^n \theta^T \sigma^{(n,m)} \theta \to \theta^T \sigma \theta$ as $n \to \infty$.

Since each $\sigma^{(n,m)}$ is positive semidefinite, $\theta^T \sigma \theta \ge 0$. Using these conditions and hypothesis (ii), we may apply Theorem 5.2 to conclude that

$$Y_{n,1} + \cdots + Y_{n,n} = \theta \cdot S^{(n)} \Rightarrow \sqrt{\theta^T \sigma \theta}\, \chi.$$

Now, $E[\sqrt{\theta^T \sigma \theta}\, \chi] = 0 = E[\theta \cdot N]$ and $E[\theta^T \sigma \theta \chi^2] = \theta^T \sigma \theta = E|\theta \cdot N|^2$, so $\sqrt{\theta^T \sigma \theta}\, \chi = \theta \cdot N$ in distribution and $\theta \cdot S^{(n)} \Rightarrow \theta \cdot N$. □

**Theorem 5.11.** (Poisson Convergence)[2]
*For each n, let $X_{n,m}$, $1 \le m \le n$ be independent nonnegative integer valued random variables and set $p_{n,m} = P(X_{n,m} = 1)$, $\varepsilon_{n,m} = P(X_{n,m} \ge 2)$. If*

*(i)* $\sum_{m=1}^n p_{n,m} \to \lambda \in (0, \infty)$ *as* $n \to \infty$,

*(ii)* $\max(p_{n,1}, p_{n,2}, \ldots, p_{n,m}) \to 0$ *as* $n \to \infty$, *and*

*(iii)* $\sum_{m=1}^n \varepsilon_{n,m} \to 0$ *as* $n \to \infty$,

*then* $S_n = X_{n,1} + \cdots + X_{n,n} \Rightarrow Z$, *where* $Z$ *is* Poisson($\lambda$).

**Theorem 5.12.** (Convergence to Stable Laws)[2]
*Let $X_1, X_2, \ldots$ be iid. Let $S_n = X_1 + \cdots + X_n$, $a_n = \inf\{x : P(|X_1| > x) \le n^{-1}\}$, and $b_n = nE[X_1 1_{\{|X_1| \le a_n\}}]$. Define*

$$w_\alpha(t) = \begin{cases} \tan(\pi\alpha/2) & \text{if } \alpha \ne 1, \\ (2/\pi) \log |t| & \text{if } \alpha = 1. \end{cases}$$

*If*

*(i)* $P(X_1 > x)/P(|X_1| > x) \to \theta \in [0, 1]$ *as* $x \to \infty$, *and*

*(ii)* $P(|X_1| > x) = x^{-\alpha} L(x)$, *where* $0 < \alpha < 2$ *and* $L$ *is slowly varying,*

*then* $(S_n - b_n)/a_n \Rightarrow Y$, *where* $Y$ *has a stable law and satisfies*

$$Ee^{itY} = \exp\{itc - b|t|^\alpha (1 + i(2\theta - 1) \operatorname{sgn}(t) w_\alpha(t))\},$$

*for some constants* $b$ *and* $c$.

**Theorem 5.13.** (Martingale Central Limit Theorem)[4]
*Let $X_1, X_2, \ldots$ be random variables and let $\mathcal{F}_1, \mathcal{F}_2, \ldots$ be σ-algebras with $\mathcal{F}_n \subset \mathcal{F}_{n+1}$ for all n. If $E(X_n|\mathcal{F}_{n-1}) = 0$, $E(X_n^2|\mathcal{F}_{n-1}) = 1$, and $E(|X_n|^3|\mathcal{F}_{n-1}) \le K < \infty$ for all n, then $(X_1 + \cdots + +X_n)/n^{1/2} \Rightarrow \chi$.*

For the next theorem, we need to start with a lemma.

**Lemma 5.14.** *Let $Z^{(n)}, Z$ be $\mathbb{R}^d$-valued random vectors such that $Z^{(n)} \Rightarrow Z$. Let $a^{(n)}, a \in \mathbb{R}^d$ satisfy $a^{(n)} \to a$ and define the set $A = \{x \in \mathbb{R}^d : x \geq a\}$. If $P(Z \in \partial A) = 0$, then $P(Z^{(n)} \geq a^{(n)}) \to P(Z \geq a)$.*

**Proof.** Let $A_n = \{x \in \mathbb{R}^d : x \geq a^{(n)}\}$. By Theorem 4.1(v) and the triangle inequality, it suffices to show that $|P(Z^{(n)} \in A_n) - P(Z^{(n)} \in A)| \leq P(Z^{(n)} \in A_n \Delta A) \to 0$. (Here, "$\Delta$" denotes the symmetric difference: $A \Delta B = (A \setminus B) \cup (B \setminus A)$.)

Fix $\delta > 0$ and let $B_\varepsilon$ denote the set of all $x \in \mathbb{R}^d$ that satisfy $x_i \geq a_i - \varepsilon$ for all $i$, and for which there exists $j$ such that $x_j \leq a_j + \varepsilon$. Since $B_\varepsilon$ is a decreasing family of sets as $\varepsilon \downarrow 0$ with $\cap_\varepsilon B_\varepsilon \subset \partial A$, we may choose $\varepsilon$ sufficiently small so that $P(Z \in B_\varepsilon) < \delta$.

Since $a^{(n)} \to a$, we have $A_n \Delta A \subset B_\varepsilon$ for $n$ sufficiently large. Hence, $P(Z^{(n)} \in A_n \Delta A) \leq P(Z^{(n)} \in B_\varepsilon)$. Since $B_\varepsilon$ is closed, Theorem 4.1(iii) gives that $\limsup_{n\to\infty} P(Z^{(n)} \in B_\varepsilon) \leq P(Z \in B_\varepsilon)$. Thus, for $n$ sufficiently large, $P(Z^{(n)} \in B_\varepsilon) \leq P(Z \in B_\varepsilon) + \delta < 2\delta$. Since $\delta$ was arbitrary, this completes the proof. □

**Theorem 5.15.** (Multi-Dimensional Quantile Central Limit Theorem)
*Let $X^{(n)} = (X_1^{(n)}, \dots, X_d^{(n)})$, $n \in \mathbb{N}$, be iid random vectors and let $F_j(x) = P(X_j^{(1)} \leq x)$. Fix $\alpha \in (0,1)^d$ and suppose there exists $q \in \mathbb{R}^d$ such that $F_j(q_j) = \alpha_j$ and $F_j'(q_j) > 0$ for all $j$. Let $M_j^{(n)}$ be the $\lfloor \alpha_j n \rfloor$-th order statistic of the quantities $X_j^{(1)}, \dots, X_j^{(n)}$ and $M^{(n)} = (M_1^{(n)}, \dots, M_d^{(n)})$. If $G_{ij}(x,y) = P(X_i^{(1)} \leq x, X_j^{(1)} \leq y)$ is continuous at $(q_i, q_j)$ for all $i, j$, then $\sqrt{n}(M_n - q) \Rightarrow N$, where $N$ is multinormal with mean $0$ and covariance $\sigma$, given by*

$$\sigma_{ij} = \frac{\rho_{ij}}{F_i'(q_i)F_j'(q_j)},$$

*with $\rho_{ij} = G_{ij}(q_i, q_j) - \alpha_i \alpha_j$.*

**Proof.** Fix $x \in \mathbb{R}^d$ and for each $n, m \in \mathbb{N}$, $1 \leq m \leq n$, define the random vector $Y^{(n,m)} \in \mathbb{R}^d$ by

$$Y_j^{(n,m)} = \frac{1}{\sqrt{n}} \left( 1_{\{X_j^{(m)} \leq x_j/\sqrt{n} + q_j\}} - p_j^{(n)} \right),$$

where $p_j^{(n)} = F_j(x_j/\sqrt{n} + q_j)$. Then for each $n \in \mathbb{N}$,

(a) $Y^{(n,m)}$, $1 \leq m \leq n$, are independent,

(b) $EY^{(n,m)} = 0$,

(c) $\sum_{m=1}^n E[Y_i^{(n,m)} Y_j^{(n,m)}] \to \rho_{ij}$ as $n \to \infty$, and

(d) for each $\theta \in \mathbb{R}^d$ and $\varepsilon > 0$, $\sum_{m=1}^n E[|\theta \cdot Y^{(n,m)}|^2 1_{\{|\theta \cdot Y^{(n,m)}| > \varepsilon)\}}] \to 0$ as $n \to \infty$.

Part (c) follows since

$$\begin{aligned}\sum_{m=1}^n E[Y_i^{(n,m)} Y_j^{(n,m)}] &= \frac{1}{n} \sum_{m=1}^n \left[ P\left( X_i^{(m)} \leq \frac{x_i}{\sqrt{n}} + q_i, X_j^{(m)} \leq \frac{x_j}{\sqrt{n}} + q_j \right) - p_i^{(n)} p_j^{(n)} \right] \\ &= P\left( X_i^{(1)} \leq \frac{x_i}{\sqrt{n}} + q_i, X_j^{(1)} \leq \frac{x_j}{\sqrt{n}} + q_j \right) - p_i^{(n)} p_j^{(n)},\end{aligned}$$

and part (d) follows since $|\theta \cdot Y^{(n,m)}| \le \max(|\theta_1|, \ldots, |\theta_d|)/\sqrt{n}$, and therefore $P(|\theta \cdot Y^{(n,m)}| > \varepsilon) = 0$ for sufficiently large $n$.

Thus, by Theorem 5.10, $S^{(n)} = Y^{(n,1)} + \cdots + Y^{(n,n)} \Rightarrow \widetilde{N}$, where $\widetilde{N}$ is multinormal with mean 0 and covariance $\rho$. Now,

$$\begin{aligned}
\sqrt{n}(M^{(n)} - q) \le x \quad &\text{iff} \quad M_j^{(n)} \le x_j/\sqrt{n} + q_j \text{ for all } j \\
&\text{iff} \quad \sum_{m=1}^{n} 1_{\{X_j^{(m)} \le x_j/\sqrt{n}+q_j\}} \ge \lfloor \alpha_j n \rfloor \text{ for all } j \\
&\text{iff} \quad \frac{1}{\sqrt{n}} \sum_{m=1}^{n} \left( 1_{\{X_j^{(m)} \le x_j/\sqrt{n}+q_j\}} - p_j^{(n)} \right) \ge \frac{\lfloor \alpha_j n \rfloor - n p_j^{(n)}}{\sqrt{n}} \text{ for all } j.
\end{aligned}$$

Thus, if $a^{(n)} \in \mathbb{R}^d$ is defined by $a_j^{(n)} = (\lfloor \alpha_j n \rfloor - n p_j^{(n)})/\sqrt{n}$, then $P(\sqrt{n}(M^{(n)} - q) \le x) = P(S^{(n)} \ge a^{(n)})$. Note that

$$a_j^{(n)} = \frac{\lfloor \alpha_j n \rfloor - \alpha_j n}{\sqrt{n}} + \sqrt{n}(\alpha_j - p_j^{(n)}) = \frac{\lfloor \alpha_j n \rfloor - \alpha_j n}{\sqrt{n}} + \frac{F_j(q_j) - F_j(x_j/\sqrt{n} + q_j)}{1/\sqrt{n}},$$

so that $a^{(n)} \to a \in \mathbb{R}^d$, where $a_j = -x_j F_j'(q_j)$. Therefore, by Lemma 5.14,

$$\begin{aligned}
P(\sqrt{n}(M^{(n)} - q) \le x) &\to P(\widetilde{N} \ge a) \\
&= P(\widetilde{N} \le -a) \\
&= P(N \le x),
\end{aligned}$$

where $N$ is the random vector defined by $N_j = \widetilde{N}_j / F_j'(q_j)$.

We now have $\sqrt{n}(M^{(n)} - q) \Rightarrow N$, $N$ is multinormal with mean 0, and

$$E[N_i N_j] = \frac{1}{F_i'(q_i) F_j'(q_j)} E[\widetilde{N}_i \widetilde{N}_j] = \frac{\rho_{ij}}{F_i'(q_i) F_j'(q_j)} = \sigma_{ij},$$

which completes the proof. □

**Corollary 5.16.** (Median Central Limit Theorem)
*Let $X_1, X_2, \ldots$ be iid, $F(x) = P(X_1 \le x)$, and $M_n = \operatorname{med}(X_1, \ldots, X_n)$. If $F(0) = 1/2$ and $F'(0) > 0$, then $\sqrt{n} M_n \Rightarrow (2F'(0))^{-1} \chi$.*

**Corollary 5.17.** (Median of Multinormal Random Vectors)
*If $X^{(1)}, X^{(2)}, \ldots$ are iid, mean 0, multinormal $\mathbb{R}^d$-valued random vectors with covariance $\sigma$ and $M^{(n)} = \operatorname{med}(X^{(1)}, \ldots, X^{(n)})$, then $\sqrt{n} M^{(n)} \Rightarrow Z$, where $Z$ is multinormal with mean 0 and covariance*

$$\tau_{ij} = \sqrt{\sigma_{ii}\sigma_{jj}} \sin^{-1}\left( \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}\sigma_{jj}}} \right),$$

*where $\sin^{-1}(\cdot)$ takes values in $[-\pi/2, \pi/2]$.*

**Proof.** By Theorem 5.15, $\sqrt{n}M^{(n)} \Rightarrow Z$, where $Z$ is multinormal with mean 0 and covariance

$$\tau_{ij} = \frac{\rho_{ij}}{F_i'(0)F_j'(0)},$$

where $\rho_{ij} = P(X_i^{(1)} \leq 0, X_j^{(1)} \leq 0) - 1/4$ and

$$F_j(x) = P(X_j^{(1)} \leq x) = \frac{1}{\sqrt{2\pi\sigma_{jj}}}\int_{-\infty}^{x} e^{-t^2/2\sigma_{jj}}\,dt.$$

Since $F_j'(0) = (2\pi\sigma_{jj})^{-1/2}$, it remains only to show that

$$\rho_{ij} = \frac{1}{2\pi}\sin^{-1}\left(\frac{\sigma_{ij}}{\sqrt{\sigma_{ii}\sigma_{jj}}}\right).$$

Let $X = X_i^{(1)}$, $Y = X_j^{(1)}$ and define

$$\begin{aligned} a^{\pm} &= 1 \pm \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}\sigma_{jj}}} \\ \widetilde{X}^{\pm} &= \frac{1}{\sqrt{2a^{\pm}}}\left(\frac{1}{\sqrt{\sigma_{ii}}}X \pm \frac{1}{\sqrt{\sigma_{jj}}}Y\right), \end{aligned}$$

so that $\widetilde{X}^{+}$, $\widetilde{X}^{-}$ are independent standard normals. Since

$$\begin{aligned} X &= \frac{\sqrt{\sigma_{ii}}}{2}\left(\sqrt{2a^{+}}\widetilde{X}^{+} + \sqrt{2a^{-}}\widetilde{X}^{-}\right) \\ Y &= \frac{\sqrt{\sigma_{jj}}}{2}\left(\sqrt{2a^{+}}\widetilde{X}^{+} - \sqrt{2a^{-}}\widetilde{X}^{-}\right), \end{aligned}$$

we have that $X \leq 0$ and $Y \leq 0$ if and only if $(\widetilde{X}^{+}, \widetilde{X}^{-})$ lies in a sector whose angle $\theta$ satisfies $0 \leq \theta \leq \pi$ and

$$\cos\theta = -\frac{2a^{+} - 2a^{-}}{2a^{+} + 2a^{-}} = -\frac{\sigma_{ij}}{\sqrt{\sigma_{ii}\sigma_{jj}}}.$$

Thus,

$$P(X \leq 0, Y \leq 0) = \frac{\theta}{2\pi} = \frac{1}{2\pi}\cos^{-1}\left(-\frac{\sigma_{ij}}{\sqrt{\sigma_{ii}\sigma_{jj}}}\right) = \frac{1}{4} + \frac{1}{2\pi}\sin^{-1}\left(\frac{\sigma_{ij}}{\sqrt{\sigma_{ii}\sigma_{jj}}}\right),$$

where $\sin^{-1}(\cdot)$ takes values in $[-\pi/2, \pi/2]$. □

# References


[1] Patrick Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999. A Wiley-Interscience Publication.

[2] Richard Durrett. *Probability: theory and examples.* Duxbury Press, Belmont, CA, second edition, 1996.

[3] Nobuyuki Ikeda and Shinzo Watanabe. *Stochastic differential equations and diffusion processes*, volume 24 of *North-Holland Mathematical Library.* North-Holland Publishing Co., Amsterdam, second edition, 1989.

[4] Jean Jacod and Philip Protter. *Probability essentials.* Universitext. Springer-Verlag, Berlin, second edition, 2003.

[5] R.-D. Reiss. *Approximate distributions of order statistics.* Springer Series in Statistics. Springer-Verlag, New York, 1989. With applications to nonparametric statistics.

[6] A. N. Shiryaev. *Probability*, volume 95 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, second edition, 1996. Translated from the first (1980) Russian edition by R. P. Boas.

[7] Daniel W. Stroock. *Probability theory, an analytic view.* Cambridge University Press, Cambridge, 1993.