Chapter 8

Laws of large numbers

We first discuss the notion of "tail equivalence" of a sequence of random variables. Here, the Borel-Cantelli Lemma is very useful. Recall that it says that if $\{E_n\}_{n\in\mathbb{N}}$ is a sequence of events with $\sum_{n=1}^{\infty} P(E_n) < \infty$, then $P\left(\limsup_{n\to\infty} E_n\right) = 0$.

Definition 8.1. *Two sequences of random variables $\{X_n\}_{n\in\mathbb{N}}$ and $\{Y_n\}_{n\in\mathbb{N}}$ are tail equivalent if*

$$\sum_{n=1}^{\infty} P\left(\{\omega : X_n(\omega) \neq Y_n(\omega)\}\right) = \sum_{n=1}^{\infty} P(\{\omega : X_n(\omega) - Y_n(\omega) \neq 0\}) = \sum_{n=1}^{\infty} P(A_n) < \infty,$$

where $A_n = \{\omega : X_n(\omega) - Y_n(\omega) \neq 0\}$.

Theorem 8.1. *Suppose $\{X_n\}_{n\in\mathbb{N}}$ and $\{Y_n\}_{n\in\mathbb{N}}$ are tail equivalent. Then,*

1. *$\sum_{n=1}^{\infty}(X_n - Y_n)$ converges almost surely,*
2. *$\sum_{n=1}^{\infty} X_n$ converges as $\iff$ $\sum_{n=1}^{\infty} Y_n$ converges as,*
3. *If there exists $a_n \to \infty$ and a random variable X such that $\frac{1}{a_n}\sum_{j=1}^{n} X_j \stackrel{as}{\to} X$, then $\frac{1}{a_n}\sum_{j=1}^{n} Y_j \stackrel{as}{\to} X$.*

Proof. 1. By tail equivalence and the Borel-Cantelli Lemma $P\left(\limsup_{n\to\infty} A_n\right) = 0$. Now, recall that $\limsup_{n\to\infty} A_n = \cap_{n=1}^{\infty}\cup_{m=n}^{\infty} A_m = \cap_{n=1}^{\infty} C_n$, where $C_n := \cup_{m=n}^{\infty} A_m$. Consequently,

$$\left(\limsup_{n\to\infty} A_n\right)^c = (\cap_{n=1}^{\infty} C_n)^c = \cup_{n=1}^{\infty} C_n^c = \cup_{n=1}^{\infty}(\cup_{m=n}^{\infty} A_m)^c = \cup_{n=1}^{\infty} \cap_{m=n}^{\infty} A_m^c = \liminf_{n\to\infty} A_n^c.$$

Thus,

$$P\left(\liminf_{n\to\infty}\{\omega : X_n(\omega) = Y_n(\omega)\}\right) = P\left(\liminf_{n\to\infty} A_n^c\right) = 1 - P\left(\limsup_{n\to\infty} A_n\right) = 1.$$

Since $\liminf_{n\to\infty} A_n^c = \{\omega : \sum_{n=1}^{\infty} I_{A_n}(\omega) < \infty\}$, $P\left(\{\omega : \sum_{n=1}^{\infty} I_{A_n}(\omega) < \infty\}\right) = 1$. Hence, there exists a set of ω's which occurs with probability 1, and in this set $X_n(\omega) = Y_n(\omega)$ for all but finitely many n. That is, for $\omega \in \{\omega : \sum_{n=1}^{\infty} I_{A_n}(\omega) < \infty\}$ there are only finitely many n for which $I_{\{X_n(\omega)\neq Y_n(\omega\}}(\omega) = 1$. That is, there exists $N(\omega)$ such that for all $n > N(\omega), I_{\{X_n(\omega)\neq Y_n(\omega)\}}(\omega) = 0$. Hence, in this same set,

$$\sum_{n=1}^{\infty} X_n(\omega) - \sum_{n=1}^{\infty} Y_n(\omega) = \sum_{n=1}^{N(\omega)} (X_n(\omega) - Y_n(\omega)) < \infty \text{ almost surely .}$$

2. Note that

$$\begin{aligned}\sum_{n=1}^{\infty} Y_n(\omega) &= \sum_{n=1}^{\infty} X_n(\omega) + \sum_{n=1}^{\infty} Y_n(\omega) - \sum_{n=1}^{\infty} X_n(\omega) \\ &= \sum_{n=1}^{\infty} X_n(\omega) - \sum_{n=1}^{\infty} (X_n(\omega) - Y_n(\omega)).\end{aligned}$$

If $\sum_{n=1}^{\infty} X_n(\omega)$ converges *as* and X_n and Y_n are tail equivalent, then both terms on the right side of the equality converge *as*, hence $\sum_{n=1}^{\infty} Y_n(\omega) < \infty$ *as*. Similarly, writing

$$\begin{aligned}\sum_{n=1}^{\infty} X_n(\omega) &= \sum_{n=1}^{\infty} X_n(\omega) + \sum_{n=1}^{\infty} Y_n(\omega) - \sum_{n=1}^{\infty} Y_n(\omega) \\ &= \sum_{n=1}^{\infty} Y_n(\omega) - \sum_{n=1}^{\infty} (Y_n(\omega) - X_n(\omega)).\end{aligned}$$

we conclude $\sum_{n=1}^{\infty} X_n(\omega) < \infty$ *as*.

3. Write

$$\begin{aligned}\frac{1}{a_n}\sum_{j=1}^{n} Y_j(\omega) &= \frac{1}{a_n}\sum_{j=1}^{n} (Y_j(\omega) - X_j(\omega) + X_j(\omega)) \\ &= \frac{1}{a_n}\sum_{j=1}^{n} (Y_j(\omega) - X_j(\omega)) + \frac{1}{a_n}\sum_{j=1}^{n} X_j(\omega) \\ &= \frac{1}{a_n}\sum_{j=1}^{N-1} (Y_j(\omega) - X_j(\omega)) + \frac{1}{a_n}\sum_{j=N}^{n} (Y_j(\omega) - X_j(\omega)) + \frac{1}{a_n}\sum_{j=1}^{n} X_j(\omega).\end{aligned}$$

As $n \to \infty$ the last term converges as to $X(\omega)$ by assumption. The second term converges to zero since $Y_j(\omega)$ and $X_j(\omega)$ are tail equivalent (and by 1), and the first term goes to 0 as $a_n \to \infty$. Hence, $\frac{1}{a_n}\sum_{j=1}^n Y_j(\omega) \overset{as}{\to} X(\omega)$.

■

The following definition and associated notation will be useful.

Definition 8.2. *Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of random variables defined on $(\Omega, \mathcal{F}, P)$ and $\{s_n\}_{n\in\mathbb{N}}$ be a sequence in $(0,\infty)$. We write,*

1. *$X_n = O_p(s_n)$ if for all $\epsilon > 0$ and $n \in \mathbb{N}$, there exists $B_\epsilon > 0$ such that*

$$P\left(\left\{\omega : \frac{|X_n(\omega)|}{s_n} > B_\epsilon\right\}\right) < \epsilon$$

2. *$X_n = o_p(s_n)$ if $\frac{X_n}{s_n} \overset{p}{\to} 0$.*

Theorem 8.2. *(General Law of Large Numbers) Suppose $\{X_n\}_{n\in\mathbb{N}}$ is a sequence of independent random variables defined on $(\Omega, \mathcal{F}, P)$ and $S_n = \sum_{j=1}^n X_j$. If*

1. *$\sum_{j=1}^n P(\{\omega : |X_j(\omega)| > n\}) \to 0$ as $n \to \infty$ and*

2. *$\frac{1}{n^2}\sum_{j=1}^n E(X_j^2 I_{\{\omega:|X_j|\le n\}}) \to 0$ as $n \to \infty$,*

then $\frac{S_n}{n} - \frac{1}{n}\sum_{j=1}^n E(X_j I_{\{\omega:|X_j|\le n\}}) \overset{p}{\to} 0$.

Proof. Let $T_{n,j}(\omega) = X_j(\omega) I_{\{\omega:|X_j|\le n\}}$ and $S_n'(\omega) = \sum_{j=1}^n T_{n,j}(\omega)$. Note that $\{\omega : X_j(\omega) \neq T_{n,j}(\omega)\} = \{\omega : |X_j(\omega)| > n\}$ and by assumption $\sum_{j=1}^n P(\{\omega : T_{n,j}(\omega) \neq X_j(\omega)\}) \to 0$ as $n \to \infty$. Note also that

$$|S_n(\omega) - S_n'(\omega)| = \left|\sum_{j=1}^n X_j(\omega) - \sum_{j=1}^n T_{n,j}(\omega)\right| \le \sum_{j=1}^n |X_j(\omega) - T_{n,j}(\omega)|.$$

Thus, for all $\epsilon > 0$,

$$\{\omega : |S_n(\omega) - S'_n(\omega)| > \epsilon\} \subset \left\{\omega : \sum_{j=1}^{n} |X_j(\omega) - T_{n,j}(\omega)| > \epsilon\right\}$$
$$\subset \bigcup_{j=1}^{n} \{\omega : |X_j(\omega) - T_{n,j}(\omega)| > \epsilon/n\}.$$

Consequently,

$$P(\{\omega : |S_n(\omega) - S'_n(\omega)| > \epsilon\}) \leq \sum_{j=1}^{n} P\left(\{\omega : |X_j(\omega) - T_{n,j}(\omega)| > \epsilon/n\}\right)$$
$$\leq \sum_{j=1}^{n} P\left(\{\omega : |X_j| > n\}\right).$$

Taking limits on both sides as $n \to \infty$, we have that $S_n - S'_n \xrightarrow{p} 0$ since by assumption 1 $\sum_{j=1}^{n} P\left(\{\omega : |X_j| > n\}\right) \to 0$.

Now, since $\{X_n\}_{n\in\mathbb{N}}$ is an independent sequence $E\left((T_{n,k} - E(T_{n,k}))(T_{n,l} - E(T_{n,l}))\right) = 0$ and consequently $V(S'_n) = \sum_{j=1}^{n} V(T_{n,j}) \leq \sum_{j=1}^{n} E(T^2_{n,j})$. Note also that for given n

$$E(T^2_{n,j}) = \int_\Omega X_j^2 I_{\{\omega:|X_j|\leq n\}} dP \leq n^2 \int_\Omega dP = n^2.$$

Consequently, since $V(S'_n)$ exists for every n, by Chebyshev's Inequality (Remark 5.1),

$$P\left(\left\{\omega : \left|\frac{S'_n - E(S'_n)}{n}\right| > \epsilon\right\}\right) \leq \frac{V(S'_n)}{n^2\epsilon^2} \leq \frac{1}{n^2\epsilon^2} \sum_{j=1}^{n} E\left(X_j^2 I_{\{\omega:|X_j|\leq n\}}\right).$$

Taking limits on both sides as $n \to \infty$ and by the assumption that $\frac{1}{n^2}\sum_{j=1}^{n} E\left(X_j^2 I_{\{\omega:|X_j|\leq n\}}\right) \to 0$, we have $\frac{S'_n}{n} - \frac{E(S'_n)}{n} \xrightarrow{p} 0$. Now, since

$$\frac{S_n}{n} - E\left(\frac{S'_n}{n}\right) = \frac{S_n}{n} - \frac{S'_n}{n} + \frac{S'_n}{n} - E\left(\frac{S'_n}{n}\right)$$

we can immediately conclude that $\frac{S_n}{n} - E\left(\frac{S'_n}{n}\right) = o_p(1)$. Finally, from the definition of S'_n we have that $\frac{S_n}{n} - \frac{1}{n}\sum_{j=1}^{n} E(X_j I_{\{\omega:|X_j|\leq n\}}) = o_p(1)$. ■

We note $E(X_j) < \infty$ or $E(X_j^2) < \infty$ are not required for Theorem 8.2. The following are examples of how Theorem 8.2 can be used.

Example 8.1. *Let* $\{X_n\}_{n\in\mathbb{N}}$ *be an independent and identically distributed sequence of random variables with* $E(X_n) = \mu$, $E(X_n^2) \leq C < \infty$. *Then, we verify condition 1 by noting that the identical distribution assumption and Markov's Inequality give*

$$\sum_{j=1}^{n} P(|X_j| > n) = nP(|X_1| > n) \leq n\frac{E(X_1^2)}{n^2} = \frac{1}{n}E(X_1^2) \leq \frac{C}{n}.$$

Taking limits on both sides as $n \to \infty$ *gives* $\lim\limits_{n\to\infty} \sum_{j=1}^{n} P(|X_j| > n) = 0$. *For condition 2, note that by the identical distribution assumption*

$$\frac{1}{n^2}\sum_{j=1}^{n} E(X_j^2 I_{\{|X_j|\leq n\}}) = \frac{1}{n}E(X_1^2 I_{\{|X_1|\leq n\}}) \leq \frac{1}{n}E(X_1^2) \leq \frac{C}{n}.$$

Again, taking limits on both sides as $n \to \infty$ *gives* $\lim\limits_{n\to\infty} \frac{1}{n^2}\sum_{j=1}^{n} E(X_j^2 I_{\{|X_j|\leq n\}}) = 0$. *Finally, observe that*

$$\frac{\sum_{j=1}^{n} E(X_j I_{\{|X_j|\leq n\}})}{n} = E(X_1 I_{\{|X_1|\leq n\}}) \to E(X_1) = \mu$$

as $n \to \infty$ *by Lebesgue's dominated convergence theorem. Thus,* $\frac{1}{n}S_n \xrightarrow{p} \mu$.

Example 8.2. *Let* $\{X_n\}_{n\in\mathbb{N}}$ *be an independent and identically distributed sequence with* $E(|X_1|) \leq C < \infty$ *and let* $E(X_1) = \mu$. *For condition 1, note that*

$$\sum_{j=1}^{n} P(|X_j| > n) = nP(|X_1| > n) = E(nI_{\{\omega:|X_1|>n\}}).$$

But since $nI_{\{\omega:|X_1|>n\}} \leq |X_1|I_{\{\omega:|X_1|>n\}}$, *we have that*

$$\sum_{j=1}^{n} P(|X_j| > n) \leq E(|X_1|I_{\{\omega:|X_1|>n\}})$$

Consequently, $\lim\limits_{n\to\infty} \sum_{j=1}^{n} P(|X_j| > n) \leq \lim\limits_{n\to\infty} E(|X_1|I_{\{\omega:|X_1|>n\}})$. *And since* $E(|X_1|) < C$, $\lim\limits_{n\to\infty} E(|X_1|I_{\{\omega:|X_1|>n\}}) = 0$.

For condition 2, note that by the identical distribution assumption

$$\begin{aligned}\frac{1}{n^2}\sum_{j=1}^{n} E\left(X_j^2 I_{\{\omega:|X_j|\leq n\}}\right) &= \frac{1}{n}E\left(X_1^2 I_{\{\omega:|X_1|\leq n\}}\right) \\ &= \frac{1}{n}\left(E\left(X_1^2 I_{\{\omega:|X_1|\leq \epsilon\sqrt{n}\}}\right) + E\left(X_1^2 I_{\{\omega:\epsilon\sqrt{n}\leq |X_1|\leq n\}}\right)\right) \text{ for any } \epsilon \in (0,1)\end{aligned}$$

Since $E(X_1^2 I_{\{\omega:|X_1|\leq \epsilon\sqrt{n}\}}) = \int_\Omega X_1^2 I_{\{\omega:|X_1|\leq \epsilon\sqrt{n}\}} dP \leq n\epsilon^2 \int_\Omega dP = n\epsilon^2$, *we have*

$$\frac{1}{n^2}\sum_{j=1}^{n} E\left(X_j^2 I_{\{\omega:|X_j|\leq n\}}\right) \leq \epsilon^2 + \frac{1}{n}E\left(|X_1||X_1|I_{\{\omega:\epsilon\sqrt{n}\leq|X_1|\leq n\}}\right) \leq \epsilon^2 + \frac{1}{n}E(n|X_1|I_{\{\omega:\epsilon\sqrt{n}\leq|X_1|\leq n\}}))$$
$$\leq \epsilon^2 + E\left(|X_1|I_{\{\omega:\epsilon\sqrt{n}\leq|X_1|\}}\right)$$

Taking limits on both sides as $n\to\infty$, *and noting that* $E(|X_1|) < C$, *we have that*

$$\lim_{n\to\infty} E(|X_j|I_{\{\omega:\epsilon\sqrt{n}\leq|X_j|\}}) = 0.$$

And, since ϵ *can be made arbitrarily small,* $\lim_{n\to\infty}\frac{1}{n^2}\sum_{j=1}^n E(X_j^2 I_{\{\omega:|X_j|\leq n\}}) = 0$. *Consequently,* $\frac{S_n}{n} - E(X_1 I_{\{\omega:|X_1|\leq n\}}) \xrightarrow{p} 0$. *Lastly, note that*

$$\lim_{n\to\infty}\left(\int_\Omega X_1 dP - \int_\Omega X_1 I_{\{|X_1|\leq n\}} dP\right) = \int_\Omega X_1 dP - \lim_{n\to\infty}\int_\Omega X_1 I_{\{|X_1|\leq n\}} dP = E(X_1) - E(X_1) = 0$$

by the previous example. Hence,

$$\frac{S_n}{n} - E(X_1) = \frac{S_n}{n} + E(X_1 I_{\{|X_1|\leq n\}}) - E(X_1 I_{\{|X_1|\leq n\}}) - E(X_1) = o_p(1) + o(1) = o_p(1).$$

Example 8.3. *Suppose* $\{X_n\}_{n\in\mathbb{N}}$ *is an independent and identically distributed sequence with* $\lim_{x\to\infty} xP(|X_1| > x) = 0$. *For condition 1, given the identically distributed assumption, we have*

$\sum_{j=1}^n P(|X_j|>n) = nP(|X_j|>n) \to 0$ *by assumption. For condition 2, note that*

$$\begin{aligned}
\frac{1}{n^2}\sum_{j=1}^n E(X_j^2 I_{\{\omega:|X_j|\le n\}}) &= \frac{1}{n}E(X_1^2 I_{\{\omega:|X_j|\le n\}}) = \frac{1}{n}\int_{|x|\le n} x^2 dF_{X_1}(x)\\
&= \frac{2}{n}\int_{|x|\le n}\left(\int_0^{|x|} s ds\right) dF_{X_1}(x) = \frac{2}{n}\int_0^n s\left(\int_{s<|x|\le n} dF_{X_1}(x)\right) ds\\
&= \frac{2}{n}\int_0^n s(P(|X_1|\le n) - P(|X_1|<s))ds\\
&= \frac{2}{n}\int_0^n s(1-P(|X_1|>n) - 1 + P(|X_1|\ge s))ds\\
&= \frac{2}{n}\int_0^n s(P(|X_1|\ge s) - P(|X_1|>n))ds\\
&= \frac{2}{n}\int_0^n \tau(s)ds - 2P(|X_1|>n)\frac{1}{n}\int_0^n sds,\ \ \textit{where } \tau(s) = sP(|X_1|>s)\\
&= \frac{2}{n}\int_0^n \tau(s)ds - 2P(|X_1|>n)\frac{1}{n}\frac{n^2}{2}\\
&= \frac{2}{n}\int_0^n \tau(s)ds - nP(|X_1|>n) = \frac{2}{n}\int_0^n \tau(s)ds - \tau(n).
\end{aligned}$$

Since, $\tau(n)\to 0$ as $n\to\infty$, we have that for all $\epsilon>0$ there exists N_ϵ such that if $n>N_\epsilon$, $\tau(n)\le\epsilon$. Consequently,

$$\frac{1}{n}\int_0^n \tau(s)ds = \frac{1}{n}\int_0^{N_\epsilon}\tau(s)ds + \frac{1}{n}\int_{N_\epsilon}^n \tau(s)ds \le \frac{1}{n}\int_0^{N_\epsilon}\tau(s)ds + \epsilon.$$

Taking limits on both sides as $n\to\infty$ gives $\frac{1}{n}\int_0^n\tau(s)ds\to 0$. Then, $\frac{S_n}{n} - E(X_1 I_{|X_1|\le n}) \overset{p}{\to} 0$.

If $\{X_j\}_{j\in\mathbb{N}}$ with $E(X_j)<\infty$, $E(X_j^2)<M<\infty$ for all j, we have that

$$\frac{1}{n}S_n - \frac{1}{n}E(S_n) = \frac{1}{n}\sum_{j=1}^n (X_j - E(X_j)) := \frac{1}{n}\sum_{j=1}^n Z_j$$

where $E(Z_j)=0$. If $E(Z_iZ_j)=0$ for all $i\ne j$, then

$$E\left(\left(\frac{1}{n}\sum_{j=1}^n Z_j\right)^2\right) = \frac{1}{n^2}\sum_{j=1}^n E(Z_j^2) < \frac{M}{n} \to 0 \text{ as } n\to\infty.$$

Hence, $\frac{1}{n}\sum_{j=1}^n Z_j \overset{\mathcal{L}^2}{\to} 0$, and by Theorem 7.7 $\frac{1}{n}\sum_{j=1}^n Z_j \overset{p}{\to} 0$. In fact, $\frac{1}{n}\sum_{j=1}^n Z_j \overset{as}{\to} 0$ as shown in the next theorem.

Theorem 8.3. *Let $\{X_j\}_{j\in\mathbb{N}}$ with $E(X_j) < \infty$, $E(X_j^2) < M < \infty$ for all j, and assume that $E((X_j - E(X_j))(X_i - E(X_i))) = 0$ for all $i \neq j$. Then, letting $Z_j = X_j - E(X_j)$ and $S_n = \sum_{j=1}^n Z_j$, we have*

$$\frac{1}{n}S_n \overset{as}{\to} 0.$$

Proof. For all $\epsilon > 0$ and by Chebyshev's Inequality

$$P(|S_n| > n\epsilon) \leq \frac{M}{n\epsilon^2}.$$

Since $\sum_{n=1}^\infty \frac{1}{n}$ diverges we can't use the Borel-Cantelli Lemma directly. However, if we consider the subsequence S_{n^2}, we have

$$\sum_{n=1}^\infty P(\{\omega : |S_{n^2}| > n^2\epsilon\}) \leq \sum_{n=1}^\infty \frac{M}{n^2\epsilon^2} < \infty.$$

Hence, $P\left(\limsup_{n\to\infty}\{\omega : |S_{n^2}(\omega)| > n^2\epsilon\}\right) = 0$ and we have $\frac{S_{n^2}}{n^2} \overset{as}{\to} 0$. Now, let

$$D_n := \max_{n^2 \leq k < (n+1)^2} |S_k - S_{n^2}|$$

and note that

$$\frac{|S_k|}{k} \leq \frac{|S_k|}{n^2} = \frac{|S_k - S_{n^2} + S_{n^2}|}{n^2} \leq \frac{|S_k - S_{n^2}|}{n^2} + \frac{|S_{n^2}|}{n^2} \leq \frac{D_n}{n^2} + \frac{|S_{n^2}|}{n^2}.$$

Now, since $P(\max_{1\leq k\leq m} |W_k| \geq \epsilon) \leq \sum_{k=1}^m P(|W_k| \geq \epsilon)$ and using Markov's Inequality

$$\begin{aligned} P(D_n \geq n^2\epsilon) &\leq \sum_{l=1}^{2n} P\left(\left|\sum_{j=1}^{l} Z_{n^2+j}\right| \geq n^2\epsilon\right) \leq \sum_{l=1}^{2n} \frac{1}{n^4\epsilon^2} E\left(\left(\sum_{j=1}^{l} Z_{n^2+j}\right)^2\right) \\ &= \sum_{l=1}^{2n} \frac{1}{n^4\epsilon^2} \sum_{j=1}^{l} E(Z_{n^2+j}^2) \leq \frac{4n^2M}{n^4\epsilon^2} = \frac{4M}{n^2\epsilon^2}. \end{aligned}$$

Then, we have $\sum_{n=1}^\infty P(D_n \geq n^2\epsilon \leq \frac{4M}{\epsilon^2}\sum_{n=1}^\infty \frac{1}{n^2} < \infty$, and by the Borel-Cantelli Lemma $\frac{D_n}{n^2} \overset{as}{\to} 0$. Since, as $n \to \infty$ we have that $k \to \infty$, $\frac{|S_k|}{k} \overset{as}{\to} 0$. ■

The following is called Markov's Law of Large Numbers.

Theorem 8.4. *(Markov's LLN) Let* $\{X_n\}_{n\in\mathbb{N}}$ *be a sequence of independent random variables with* $E(X_n) = \mu_n$. *If for some* $\delta > 0$ *we have* $\frac{1}{n}\sum_{n=1}^{\infty} \frac{E|X_n - \mu_n|^{1+\delta}}{n^{1+\delta}} < \infty$,

$$S_n - \frac{1}{n}\sum_{i=1}^{n} \mu_i \xrightarrow{p} 0.$$

Proof. Chung (1974, A Course in Probability Theory, pp. 125-126). ■