Therefore,

$$f_{W,Z}(\alpha,\beta) = \frac{f_{X,Y}(u.v)}{\det J} = \begin{cases} \frac{\exp\left(-\lambda\left(2\sqrt{\beta-\alpha}-\alpha\right)\right)}{\sqrt{\beta-\alpha}} & \beta > \alpha + (\max\{0,\alpha\})^2 \\ 0 & else \end{cases}$$

1.35 Conditional densities and expectations
(a)

$$\begin{aligned} E[XY] &= \int_0^1 \int_0^u uv(4u^2)dvdu \\ &= \int_0^1 4u^3 \left(\int_0^u vdv\right) du \\ &= \int_0^1 2u^5 du = \frac{1}{3}. \end{aligned}$$

(b)

$$f_Y(v) = \int_v^1 4u^2 \; du = \begin{cases} \frac{4}{3}(1-v^3), & 0 \le v \le 1 \\ 0, & \text{elsewhere} \end{cases}$$

(c)

$$f_{X|Y}(u|v) = \begin{cases} 0, & 0 < v < 1, \quad 0 < u < v \\ \frac{4u^2}{\frac{4}{3}(1-v^3)} = \frac{3u^2}{1-v^3}, & 0 < v < 1, \quad v < u < 1 \\ \text{undefined}, & v < 0 \text{ or } v > 1 \end{cases}$$

(d) For $0 < v < 1$, $E[X^2|Y=v] = \int_v^1 \; u^2 \; \frac{3u^2}{1-v^3} \; du = \frac{3}{5}\frac{1-v^5}{1-v^3}$

2.1 Limits and infinite sums for deterministic sequences (a) Before beginning the proof we observe that $|\cos(\theta)| \le 1$, so $|\theta(1+\cos(\theta))| \le 2|\theta|$. Now, for the proof. Given an arbitrary ϵ with $\epsilon > 0$, let $\delta = \epsilon/2$. For any θ with $|\theta - 0| \le \delta$, the following holds: $|\theta(1+\cos(\theta)) - 0| \le 2|\theta| \le 2\delta = \epsilon$. Since ϵ was arbitrary the convergence is proved.
(b) Before beginning the proof we observe that if $0 < \theta < \pi/2$, then $\cos(\theta) \ge 0$ and $\frac{1+\cos(\theta)}{\theta} \ge 1/\theta$. Now, for the proof. Given an arbitrary positive number K, let $\delta = \min\{\frac{\pi}{2}, \frac{1}{K}\}$. For any θ with $0 < \theta < \delta$, the following holds: $\frac{1+\cos(\theta)}{\theta} \ge 1/\theta \ge 1/\delta \ge K$. Since K was arbitrary the convergence is proved.
(c) The sum is by definition equal to $\lim_{N\to\infty} s_N$ where $s_N = \sum_{n=1}^N \frac{1+\sqrt{n}}{1+n^2}$. The sequence S_N is increasing in N. Note that the $n=1$ term of the sum is 1 and for any $n \ge 1$ the n^{th} term of the sum can be bounded as follows:

$$\frac{1+\sqrt{n}}{1+n^2} \le \frac{2\sqrt{n}}{n^2} = 2n^{-3/2}.$$

Therefore, comparing the partial sum with an integral, yields

$$s_N \le 1 + \sum_{n=2}^N 2n^{-3/2} \le 1 + \int_1^N 2x^{-3/2}dx = 5 - 4N^{-1/2} \le 5.$$

In summary, the sequence $(S_N : N \geq 1)$ is an increasing, bounded sequence, and it thus has a finite limit.

2.3 The reciprocal of the limit is the limit of the reciprocal Let $\epsilon > 0$. Let $\epsilon' = \min\{\frac{|x_\infty|}{2}, \frac{\epsilon x_\infty^2}{2}\}$. By the hypothesis, there exists a value of n_o so large that for all $n \geq n_o$, $|x_n - x_\infty| \leq \epsilon'$. This condition implies that $|x_n| \geq |x_\infty|/2$, because of the choice of ϵ'. Therefore, for all $n \geq n_o$,

$$\left|\frac{1}{x_n} - \frac{1}{x_\infty}\right| = \frac{|x_n - x_\infty|}{|x_n||x_\infty|} \leq \frac{2\epsilon'}{x_\infty^2} \leq \epsilon,$$

which, by definition, shows that $(1/x_n)$ converges to $1/x_\infty$.

2.5 On convergence of deterministic sequences and functions (a) Note that $x_n - \frac{8}{3} = \frac{1}{3n}$. Thus, given any $\epsilon > 0$, let $n_\epsilon = \lceil \frac{1}{3\epsilon} \rceil$. Then for any $n \geq n_\epsilon$, $|X_n - \frac{8}{3}| \leq \frac{1}{3n} \leq \frac{1}{3n_\epsilon} \leq \epsilon$. Thus, by definition, $\lim_{n\to\infty} x_n = \frac{8}{3}$.
(b) Let $\epsilon = 1/3$ and let $x_n = (2/3)^{1/n}$ for $n \geq 1$. Note that $x_n \in [0,1)$ and $f_n(x_n) = \frac{2}{3}$. Thus, there is no positive integer n such that $|f_n(x) - 0| \leq \epsilon$ for all $x \in [0,1)$. So it is impossible to select n_ϵ with the property required for uniform convergence. Therefore f_n does not converge uniformly to zero.
(c) Let $c < \sup_D f$. Then there is an $x \in D$ so that $c \leq f(x)$. Therefore, $c \leq f(x) - g(x) + g(x) \leq \sup_D |f - g| + \sup_D g$. Thus, $c < \sup_D f$ implies $c < \sup_D |f - g| + \sup_D g$. Equivalently, $\sup_D f \leq \sup_D |f - g| + \sup_D g$, or $\sup_D f - \sup_D g \leq \sup_D |f - g|$. Exchanging the roles of f and g yields $\sup_D g - \sup_D f \leq \sup_D |f - g|$. Combining yields the desired inequality, $|\sup_D f - \sup_D g| \leq \sup_D |f - g|$. As an application, suppose $f_n \to f$ uniformly on D. Then given any $\epsilon > 0$, there exists an n_ϵ so large, that $\sup_D |f_n - f| \leq \epsilon$, whenever $n \geq n_\epsilon$. But then by the inequality proved, $|\sup_D f_n - \sup_D f| \leq \sup_D |f_n - f| \leq \epsilon$, whenever $n \geq n_\epsilon$. Thus, by definition, $\sup_D f_n \to \sup_D f$ as $n \to \infty$.

2.7 On the Dirichlet criterion for convergence of a series
(a) Let $R_n = \sum_{k=0}^n a_k$. By assumption, the sequence (R_n) has a finite limit, so it is a Cauchy sequence, i.e. $\lim_{m,n\to\infty} |R_m - R_n| = 0$. Now for $n < m$,
$|S_m - S_n| = |\sum_{k=n+1}^m d_k| \leq \sum_{k=n+1}^m |d_k| \leq \sum_{k=n+1}^m La_k = L|R_m - R_n|$. Therefore,
$\lim_{m,n\to\infty} |S_m - S_n| = 0$. That is, (S_n) is also a Cauchy sequence, and hence also has a finite limit.

(b)

$$\begin{aligned}
S_n &= \sum_{k=0}^{n} A_k B_k - \sum_{k=1}^{n} A_k B_{k-1} \quad \text{since } B_{-1}=0 \\
&= \sum_{k=0}^{n} A_k B_k - \sum_{k=0}^{n-1} A_{k+1} B_k \\
&= \left(\sum_{k=0}^{n} (A_k - A_{k+1}) B_k \right) - A_{n+1} B_n \\
&= \left(\sum_{k=0}^{n} a_k B_k \right) - A_{n+1} B_n .
\end{aligned}$$

(c) Since $|a_k B_k| \leq L a_k$ for all k, the sequence of sums $\sum_{k=0}^{n} a_k B_k$ is convergent by the result of part (a). Also, $|A_{n+1} B_n| \leq L A_{n+1} \to 0$ as $n \to \infty$. Thus, by part (b), S_n has a finite limit.

2.9 Convergence of a random sequence (a) The sequence $X_n(\omega)$ is monotone nondecreasing in n for each ω. Also, by induction on n, $X_n(\omega) \leq 1$ for all n and ω. Since bounded monotone sequences have finite limits, $\lim_{n\to\infty} X_n$ exists in the a.s. sense and the limit is less than or equal to one with probability one.
(b) Since a.s. convergence of bounded sequences implies m..s. convergence, $\lim_{n\to\infty} X_n$ also exists in the m.s. sense.
(c) Since (X_n) converges a.s., it also converges in probability to the same random variable, so $Z = \lim_{n\to\infty} X_n$ a.s. It can be shown that $P\{Z = 1\} = 1$. Here is one of several proofs. Let $0 < \epsilon < 1$. Let $a_0 = 0$ and $a_k = \frac{a_{k-1}+1-\epsilon}{2}$ for $k \geq 1$. By induction, $a_k = (1-\epsilon)(1-2^{-k})$. Consider the sequence of events: $\{U_i \geq 1-\epsilon\}$ for $i \geq 1$. These events are independent and each has probability ϵ. So with probability one, for any $k \geq 1$, the probability that at least k of these events happens is one. If at least k of these events happen, then $Z \geq a_k$. So, $P\{(1-\epsilon)(1-2^{-k}) \leq Z \leq 1\} = 1$. Since ϵ can be arbitrarily close to zero and k can be arbitrarily large, it follows that $P\{Z = 1\} = 1$.
ANOTHER APPROACH is to calculate that $E[X_n|X_{n-1} = v] = v + \frac{(1-v)^2}{2}$. Thus, $E[X_n] = E[X_{n-1}] + \frac{E[(1-X_{n-1})^2]}{2} \geq E[X_{n-1}] + \frac{(1-E[X_{n-1}])^2}{2}$. Since $E[X_n] \to E[Z]$, it follows that $E[Z] \geq E[Z] + \frac{(1-E[Z])^2}{2}$. So $E[Z] = 1$. In view of the fact $P\{Z \leq 1\} = 1$, it follows that $P\{Z = 1\} = 1$.

2.11 Convergence of some sequences of random variables (a)For each fixed ω., $\frac{V(\omega)}{n} \to 0$ so $X_n(\omega) \to 1$. Thus, $X_n \to 1$ in the a.s sense, and hence also in the p. and d. senses. Since the random variables X_n are uniformly bounded (specifically, $|X_n| \leq 1$ for all n), the convergence in p. sense implies convergence in m.s. sense as well. So $X_n \to 1$ in all four senses.
(b)To begin we note that $P\{V \geq 0\} = 1$ with $P\{V > 1\} = e^{-3} > 0$. For any ω such that $V(\omega) < 1$, $Y_n(\omega) \to 0$, and for any ω such that $V(\omega) > 1$, $Y_n(\omega) \to +\infty$, so (Y_n) does not converge in the a.s. sense to a finite random variable.

Let us show Y_n does not converge in d. sense. For any $c > 0$ $\lim_{n\to\infty} F_n(c) = \lim_{n\to\infty} P\{Y_n \le c\} = P\{V < 1\} = 1 - e^{-3}$. The limit exists but the limit function F satisfies $F(c) = e^{-1}$ for all $c > 0$, so the limit is not a valid CDF. Thus, (Y_n) does not converge in the d. sense (to a finite limit random variable), and hence does not converge in any of the four senses to a finite limit random variable.
(c)For each ω fixed, $Z_n(\omega) \to e^{V(\omega)}$. So $Z_n \to e^V$ in the a.s. sense, and hence also in the p. and d. senses. Using the inequality $1+u \le e^u$ shows that $Z_n \le e^V$ for all n so that $|Z_n| \le e^V$ for all n. Note that $E[(e^V)^2] = E[e^{2V}] = \int_0^\infty e^{2u} 3e^{-3u} du = 3 < \infty$. Therefore, the sequence (Z_n) is dominated by a single random variable with finite second moment (namely, e^V), so the convergence of (Z_n) in the p. sense to e^V implies that (Z_n) converges to e^V in the m.s. sense as well. So $Z_n \to e^V$ in all four senses.
2.13 On the maximum of a random walk with negative drift (a) By the strong law of large numbers, $P\{S_n/n \to -1\} = 1$. Therefore, with probability one, $S_n/n \le 0$ for all sufficiently large n. That is, with probability one, $S_n > 0$ only finitely many times. The random variable Z, with probability one, is thus the maximum of only finitely many nonnegative numbers. So Z is finite with probability one.
(b) Suppose $P\{X_1 = c-1\} = P\{X_1 = -c-1\} = 0.5$ for a constant $c > 0$. Then X_1 has mean -1 as required. Following the hint, for $c \ge 1$, we have $E[Z] \ge E[\max\{0, X_1\}] = (c-1)/2$. Observe that $E[Z]$ can be made arbitrarily large by taking c arbitrarily large. So the answer to the question is no. (Note: More can be said about $E[Z]$ if the variance of X_1 is known. A celebrated bound of J.F.C. Kingman is that $E[Z] \le \frac{\mathrm{Var}(X_1)}{-2E[X_1]}$.)
2.15 Convergence in distribution to a nonrandom limit Suppose $P\{X = c\} = 1$ and $\lim_{n\to\infty} X_n = X$ $d.$ Let $\epsilon > 0$. It suffices to prove that
$P\{X_n - X| \le \epsilon\} \to 1$ as $n \to \infty$. Note that $P\{|X_n - X| \le \epsilon\} \ge P\{c - \epsilon < X_n \le c + \epsilon\} = F_n(c+\epsilon) - F_n(c-\epsilon)$. Since $c - \epsilon$ is a continuity point of F_X and $F_X(c-\epsilon) = 0$, it follows that $F_n(c-\epsilon) \to 0$. Similarly, $F_n(c+\epsilon) \to 1$. Thus $F_n(c+\epsilon) - F_n(c-\epsilon) \to 1$, so that $P\{|X_n - X| \le \epsilon\} \to 1$. Therefore convergence in probability holds.
Note: A slightly different approach would be to prove that for any $\epsilon > 0$, there is an n_ϵ so large that $P\{|X_n - c| \le \epsilon\} \ge 1 - \epsilon$.
2.17 Convergence of a product (a) Examine $S_n = \ln X_n$. The sequence S_n, $n \ge 1$ is the sequence of partial sums of the independent and identically distributed random variables $\ln U_k$. Observe that $E[\ln U_k] = \int_0^2 \ln(u) \frac{1}{2} du = \frac{1}{2}(x \ln x - x)|_0^2 = \ln 2 - 1 \approx -0.306$. Therefore, by the strong law of large numbers, $\lim_{n\to\infty} \frac{S_n}{n} = \ln 2 - 1$ a.s. This means that, given an $\epsilon > 0$, there is an a.s. finite random variable N_ϵ so large that $|\frac{S_n}{n} - (\ln 2 - 1)| \le \epsilon$ for all $n \ge N_\epsilon$. Equivalently,

$$\left(\frac{2(1-\epsilon)}{e}\right)^n \le X_n \le \left(\frac{2(1+\epsilon)}{e}\right)^n \quad \text{for} \quad n \ge N_\epsilon.$$

Conclude that $\lim_{n\to\infty} X_n = 0$ a.s., which implies that also $\lim_{n\to\infty} X_n = 0$ p.

and $\lim_{n\to\infty} X_n = 0$ d. It remains to check for convergence in the m.s. sense. If X_n were to converge in the m.s. sense, it would have to converge to the same random variable in probability. But X_n does not converge to zero in the mean square. One way to see that X_n does not converge to zero in the m.s. sense is to note that $E[X_n] = 1$ for all n, so $\lim_{n\to\infty} E[X_n] \neq 0$. Another way to see that X_n does not converge to zero in the m.s. sense is to observe

$$E[|X_n - 0|^2] = E[U_1^2 \dots U_n^2] = E[U_1^2] \cdots E[U_n^2] = \left(\frac{4}{3}\right)^n \not\to 0.$$

In summary, X_n converges to zero in the a.s., p., and d. senses, but does not converge in the m.s. sense.

(b) As noted in part (a), $\lim_{n\to\infty} \frac{S_n}{n} = \ln 2 - 1$ a.s., so if $\theta = -1$, then $Y_n = \frac{S_n}{n}$ converges in distribution to the constant $\ln 2 - 1$.

2.19 Sums of i.i.d. random variables, I Let X_i denote the gamblers net gain for the i^{th} play. Then $X_1, X_2, \dots, X_{100}$ are iid with

$$p_{X_i}(x) = \begin{cases} 0.5 & x = -1 \\ 0.1 & x = 0 \\ 0.4 & x = 1 \\ 0 & \text{otherwise} \end{cases}$$

Thus $E[X_i] = -0.1$, $\text{Var}(X_i) = E[X_i^2] - E[X_i]^2 = 0.9 - (0.1)^2 = .89$. Let $S = X_1 + \cdots + X_{100}$. Therefore $ES = -10$, $\text{Var}(S) = (100)(0.89) = 89$. By Chebychev's inequality

$$\begin{aligned} P(S \geq 0) &= P(S + 10 \geq 10) \\ &\leq P(|S + 10| \geq 10) \leq \frac{\text{Var}(S)}{(10)^2} = .89. \end{aligned}$$

By the central limit theorem,

$$P(S \geq 0) = P\left\{\frac{S + 10}{\sqrt{89}} \geq \frac{10}{\sqrt{89}}\right\} \approx 1 - \Phi(1.06) = 0.1446.$$

To calculate the Chernoff bound, we find $M(\theta) = \log(0.5e^{\theta} + 0.1 + 0.4e^{-\theta})$ and $\ell(0) = \exp(.4\sqrt{5} + 0.1)$, yielding the upper bound $P(S \geq 0) \leq (.4\sqrt{5} + 0.1)^{100} = 0.57187$.

It is not difficult to calculate $P(S \geq 0)$ numerically. The result is $P(S \geq 0) = 0.1572....$ Thus, in this example, the approximation based on the central limit theorem is fairly accurate, the Chernoff bound is somewhat loose, and the Chebychev inequality is very loose.

2.21 Sums of i.i.d. random variables, III (a) $\Phi_{X_{i,n}}(u) = E\left[e^{juX_{i,n}}\right] = 1 + \frac{\lambda}{n}(e^{ju} - 1)$ so $\Phi_{Y_n}(u) = \left(1 + \frac{\lambda}{n}(e^{ju} - 1)\right)^n$.

(b) Since $\lim_{n\to\infty}\left(1 + \frac{\alpha}{n}\right)^{\alpha} = e^{\alpha}$, it follows that $\lim_{n\to\infty} \Phi_{Y_n}(u) = e^{\lambda(e^{ju}-1)}$. This limit as a function of u is the characteristic function of a random variable Y with the Poisson distribution with mean λ.

(c) Thus, (Y_k) converges in distribution, and the limiting distribution is the

Poisson distribution with mean λ. There is not enough information given in the problem to determine whether Y_n converges in any of the stronger senses (p., m.s., or a.s.), because the given information only describes the distribution of Y_n for each n but gives nothing about the joint distribution of the Y_n's. Note that Y_n has a binomial distribution for each n.

2.23 On the growth of the maximum of n independent exponentials (a) Let $n \geq 2$. Clearly $F_{Z_n}(c) = 0$ for $c \leq 0$. For $c > 0$,

$$\begin{aligned} F_{Z_n}(c) &= P\{\max\{X_1, \ldots, X_n\} \leq c \ln n\} \\ &= P\{X_1 \leq c \ln n,\ X_2 \leq c \ln n,\ \ldots, X_n \leq c \ln n\} \\ &= P\{X_1 \leq c \ln n\} P\{X_2 \leq c \ln n\} \cdots P\{X_n \leq c \ln n\} \\ &= (1 - e^{-c \ln n})^n = (1 - n^{-c})^n. \end{aligned}$$

(b) Or, $F_{X_n}(c) = (1 + \frac{x_n}{n})^n$, where $x_n = -n^{1-c}$. Observe that as $n \to \infty$,

$$x_n \to \begin{cases} -\infty & c < 1 \\ -1 & c = 1 \\ 0 & c > 1, \end{cases}$$

so by Lemma 2.3.1 (and the monotonicity of the function e^x to extend to the case $x = -\infty$),

$$F_{Z_n}(c) \to \begin{cases} 0 & c < 1 \\ e^{-1} & c = 1 \\ 1 & c > 1. \end{cases}$$

Therefore, if Z_∞ is the random variable that is equal to one with probability one, then $F_{Z_n}(c) \to F_{Z_\infty}(c)$ at the continuity points (i.e. at $c \neq 1$) of F_{Z_∞}. So the sequence (Z_n) converges to one in distribution.

2.25 Limit behavior of a stochastic dynamical system Due to the persistent noise, just as for the example following Theorem 2.1.5 in the notes, the sequence does not converge to an ordinary random variables in the a.s., p., or m.s. senses. To gain some insight, imagine (or simulate on a computer) a typical sample path of the process. A typical sample sequence hovers around zero for a while, but eventually, since the Gaussian variables can be arbitrarily large, some value of X_n will cross above any fixed threshold with probability one. After that, X_n would probably converge to infinity quickly. For example, if $X_n = 3$ for some n, and if the noise were ignored from that time forward, then X would go through the sequence $9, 81, 6561, 43046721, 1853020188851841, 2.43 \times 10^{30}, \ldots$, and one suspects the noise terms would not stop the growth. This suggests that $X_n \to +\infty$ in the a.s. sense (and hence in the p. and d. senses as well. (Convergence to $+\infty$ in the m.s. sense is not well defined.) Of course, then, X_n does not converge in any sense to an ordinary random variable.

We shall follow the above intuition to *prove* that $X_n \to \infty$ a.s. If $W_{n-1} \geq 3$ for some n, then $X_n \geq 3$. Thus, the sequence X_n will eventually cross above the threshold 3. We say that X *diverges nicely from time n* if the event $E_n =$

$\{X_{n+k} \geq 3 \cdot 2^k$ for all $k \geq 0\}$ is true. Note that if $X_{n+k} \geq 3 \cdot 2^k$ and $W_{n+k} \geq -3 \cdot 2^k$, then $X_{n+k+1} \geq (3 \cdot 2^k)^2 - 3 \cdot 2^k = 3 \cdot 2^k(3 \cdot 2^k - 1) \geq 3 \cdot 2^{k+1}$. Therefore, $E_n \supset \{X_n \geq 3$ and $W_{n+k} \geq -3 \cdot 2^k$ for all $k \geq 0\}$. Thus, using a union bound and the bound $Q(u) \leq \frac{1}{2}\exp(-u^2/2)$ for $u \geq 0$:

$$\begin{aligned}
P(E_n|X_n \geq 3) &\geq P\{W_{n+k} \geq -3 \cdot 2^k \text{ for all } k \geq 0\} \\
&= 1 - P\left(\cup_{k=0}^{\infty}\{W_{n+k} \leq -3 \cdot 2^k\}\right) \\
&\geq 1 - \sum_{k=0}^{\infty} P\{W_{n+k} \leq -3 \cdot 2^k\} = 1 - \sum_{k=0}^{\infty} Q(3 \cdot 2^k \cdot \sqrt{2}) \\
&\geq 1 - \frac{1}{2}\sum_{k=0}^{\infty} \exp(-(3 \cdot 2^k)^2) \geq 1 - \frac{1}{2}\sum_{k=0}^{\infty}(e^{-9})^{k+1} \\
&= 1 - \frac{e^{-9}}{2(1-e^{-9})} \geq 0.9999.
\end{aligned}$$

The pieces are put together as follows. Let N_1 be the smallest time such that $X_{N_1} \geq 3$. Then N_1 is finite with probability one, as explained above. Then X diverges nicely from time N_1 with probability at least 0.9999. However, if X does not diverge nicely from time N_1, then there is some first time of the form $N_1 + k$ such that $X_{N_1+k} < 3 \cdot 2^k$. Note that the future of the process beyond that time has the same evolution as the original process. Let N_2 be the first time after that such that $X_{N_2} \geq 3$. Then X again has chance at least 0.9999 to diverge nicely to infinity. And so on. Thus, X will have arbitrarily many chances to diverge nicely to infinity, with each chance having probability at least 0.9999. The number of chances needed until success is a.s. finite (in fact it has the geometric distribution), so that X diverges nicely to infinity from some time, with probability one.

2.27 Convergence analysis of successive averaging (b) The means μ_n of X_n for all n are determined by the recursion $\mu_0 = 0$, $\mu_1 = 1$, and, for $n \geq 1$, $\mu_{n+1} = (\mu_n + \mu_{n-1})/2$. This second order recursion has a solution of the form $\mu_n = A\theta_1^n + B\theta_2^n$, where θ_1 and θ_2 are the solutions to the equation $\theta^2 = (1+\theta)/2$. This yields $\mu_n = \frac{2}{3}(1 - (-\frac{1}{2})^n)$.

(c) It is first proved that $\lim_{n\to\infty} D_n = 0$ *a.s.*. Note that $D_n = U_1 \cdots U_{n-1}$. Since $\ln D_n = \ln(U_1) + \cdots \ln(U_{n-1})$ and $E[\ln U_i] = \int_0^1 \ln(u)du = (x\ln x - x)|_0^1 = -1$, the strong law of large numbers implies that $\lim_{n\to\infty} \frac{\ln D_n}{n-1} = -1$ *a.s.*, which in turn implies $\lim_{n\to\infty} \ln D_n = -\infty$ *a.s.*, or equivalently, $\lim_{n\to\infty} D_n = 0$ *a.s.*, which was to be proved. By the hint, for each ω such that $D_n(\omega)$ converges to zero, the sequence $X_n(\omega)$ is a Cauchy sequence of numbers, and hence has a limit. The set of such ω has probability one, so X_n converges a.s.

2.29 Mean square convergence of a random series Let $Y_n = X_1 + \cdots + X_n$. We are interested in determining whether $\lim_{n\to\infty} Y_n$ exists in the m.s. sense. By Proposition 2.11, the m.s. limit exists if and only if the limit $\lim_{m,n\to\infty} E[Y_mY_n]$ exists and is finite. But $E[Y_mY_n] = \sum_{k=1}^{n\wedge m}\sigma_k^2$ which converges to $\sum_{k=1}^{\infty}\sigma_k^2$ as $n, m \to \infty$. Thus, (Y_n) converges in the m.s. sense if and only if $\sum_{k=1}^{\infty}\sigma_k^2 < \infty$.

2.31 A large deviation Since $E[X_1^2] = 2 > 1$, Cramér's theorem implies that $b = \ell(2)$, which we compute. For $a > 0$, $\int_{-\infty}^{\infty} e^{-ax^2}dx = \int_{-\infty}^{\infty} e^{-\frac{x^2}{2(\frac{1}{2a})}}dx = \sqrt{\frac{\pi}{a}}$, so

$$M(\theta) = \ln E[e^{\theta x^2}] = \ln \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2(\frac{1}{2}-\theta)}dx = -\frac{1}{2}\ln(1-2\theta).$$

$$\begin{aligned}
\ell(a) &= \max_{\theta}\left\{\theta a + \frac{1}{2}\ln(1-2\theta)\right\} \\
&= \frac{1}{2}\left(1 - \frac{1}{a}\right) \\
\theta^* &= \frac{1}{2}(1-\frac{1}{a}) \\
b = \ell(2) &= \frac{1}{2}(1-\ln 2) = 0.1534 \\
e^{-100b} &= 2.18 \times 10^{-7}.
\end{aligned}$$

2.33 Sums of independent Cauchy random variables
(a) $\Phi_{S_n/n^\theta}(u) = E[\exp(j\frac{u}{n^\theta}X_1)]^n = \Phi(\frac{u}{n^\theta})^n = (\exp(-|\frac{u}{n^\theta}|)^n = \exp(-|u|n^{1-\theta})$
(b) Taking $\theta = 1$ we see that $\frac{S_n}{n}$ has the same distribution as X_1 for all n. Thus, $\frac{S_n}{n}$ converges in distribution, and the limiting distribution is the standard Cauchy distribution. (This answer suggests the answers to parts (c) and (d).)
(c) Taking $\theta = 2$ yields that the characteristic function of $\frac{S_n}{n^2}$ is $\exp(-|u|/n)$ which converges to 1 for all u. But 1 is the characteristic function of a random variable that is zero with probability one. Thus. $\frac{S_n}{n^2} \to 0$ $d.$ as $n \to \infty$. A different approach to part (c) is to use part (b) to find that for any $\epsilon > 0$, $P\left\{|\frac{S_n}{n^2}| \le \epsilon\right\} = P\{|X_1| \le n\epsilon\} \to 1$ as $n \to \infty$.
(d) Taking $\theta = 1/2$ yields that the characteristic function of $\frac{S_n}{n^{1/2}}$ is $\exp(-|u|n^{1/2})$ which converges to $I_{\{u=0\}}$ for all u. But $I_{\{u=0\}}$ is not a valid characteristic function (since it is not continuous, for example) so that $\frac{S_n}{n^{1/2}}$ does not converge in distribution. A different approach to part (d) is to use part (b) to find that for any constant c,
$P\{\frac{S_n}{n^{1/2}} \le c\} = P\{X_1 \le cn^{-1/2}\} \to P\{X_1 \le 0\} \to 1/2$. But the function equal to the constant 1/2 is not a valid distribution function (it doesn't converge to 0 at $-\infty$, for example) so the same conclusion follows.

2.35 Chernoff bound for Gaussian and Poisson random variables
(a) $M_X(\theta) = \mu\theta + \frac{\theta^2\sigma^2}{2}$ and $l(a) = \max_\theta a\theta - (\mu\theta + \frac{\theta^2}{2\sigma^2}) = \frac{(a-\mu)^2}{2\sigma^2}$. Taking $a = \mu + c$ in the optimized Chernoff inequality yields $P\{X \ge E[X] + c\} \le \exp(-\frac{c^2}{2\sigma^2})$ for $c \ge 0$.
(b) The log moment generating function of Y is given by
$M_Y(\theta) = \ln \sum_{k=0}^{\infty} \frac{e^{\theta k}\lambda^k e^{-\lambda}}{k!} = \ln(e^{\lambda(e^\theta - 1)}) = \lambda(e^\theta - 1)$. Therefore,

$$l(a) = \max_\theta a\theta - e^{\lambda}(\theta - 1) = a\ln(\frac{a}{\lambda}) + \lambda - a.$$

Setting $a = \lambda + c$ in the optimized Chernoff inequality yields

$$P\{Y \geq E[Y] + c\} \leq \exp\left(-(\lambda + c)\ln\frac{\lambda + c}{\lambda} + c\right).$$

(c) Using the definition of ψ yields $\frac{c^2}{2\lambda}\psi(\frac{c}{\lambda}) = \lambda g(1 + \frac{c}{\lambda}) = (\lambda + c)\ln\frac{\lambda+c}{\lambda} - c$ as desired. For more information see Shorack and Wellner, *Empirical Processess*, 1986.

2.37 Large deviation exponent for a mixture distribution
(a) $\widetilde{l}(a) = \max_\theta \{\theta a - M_X(\theta)\}$ where

$$\begin{aligned} M_X(\theta) &= \log E[\exp(\theta X)] = \log\{fE[\exp(\theta Y)] + (1 - f)E[\exp(\theta Z)]\} \\ &= \log\{f\exp(M_Y(\theta)) + (1 - f)\exp(M_Z(\theta))\}. \end{aligned}$$

(b) View $f\exp(M_Y(\theta)) + (1 - f)\exp(M_Z(\theta))$ as an average of $\exp(M_Y(\theta))$ and $\exp(M_Z(\theta))$. The definition of concavity (or Jensen's inequality) applied to the concave function $\log u$ implies that $\log(average) \geq average(\log)$, so that $\log\{f\exp(M_Y(\theta)) + (1 - f)\exp(M_Z(\theta)) \geq fM_Y(\theta) + (1 - f)M_Z(\theta)$, where we also used the fact that $\log\exp M_Y(\theta) = M_Y(\theta)$. Therefore, $\widetilde{l}(a) \leq l(a)$ for all a. Remark: This means that $\frac{\widetilde{S}_n}{n}$ is more likely to have large deviations than $\frac{S_n}{n}$. That is reasonable, because $\frac{\widetilde{S}_n}{n}$ has randomness due not only to F_Y and F_Z, but also due to the random coin flips. This point is particularly clear in case the Y's and Z's are constant, or nearly constant, random variables.

2.39 Bernstein's inequality in various asymptotic regimes (a) The bound becomes

$$P\left\{S_n \geq \theta\sqrt{n}\right\} \leq \exp\left(-\frac{\frac{1}{2}\theta^2}{\sigma^2 + \frac{\theta L}{3\sqrt{n}}}\right).$$

(b) The bound becomes

$$P\left\{S_n \geq cn\right\} \leq \exp\left(-\frac{\frac{1}{2}c^2 n}{\sigma^2 + \frac{cL}{3}}\right).$$

(b) The bound becomes

$$P\left\{S_n \geq \alpha\right\} \leq \exp\left(-\frac{\frac{1}{2}\alpha^2}{\gamma + \frac{\alpha L}{3}}\right).$$

3.1 Rotation of a joint normal distribution yielding independence
(a) $|\mathrm{Cov}(X)| = 1$ and $\mathrm{Cov}(X)^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$ so

$$\begin{aligned} f_X(x) &= \frac{1}{2\pi}\exp\left(-\frac{1}{2}\begin{pmatrix} x_1 - 10 \\ x_2 - 5 \end{pmatrix}^T \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} x_1 - 10 \\ x_2 - 5 \end{pmatrix}\right) \\ &= \frac{1}{2\pi}\exp\left(-\frac{1}{2}\{(x_1 - 10)^2 - 2(x_1 - 10)(x_2 - 5) + 2(x_2 - 5)^2\}\right). \end{aligned}$$