Solutions of Selected Problems from *Probability Essentials, Second Edition*

Solutions to selected problems of Chapter 2

2.1 Let's first prove by induction that $\#(2^{\Omega_n}) = 2^n$ if $\Omega = \{x_1, \dots, x_n\}$. For $n = 1$ it is clear that $\#(2^{\Omega_1}) = \#(\{\emptyset, \{x_1\}\}) = 2$. Suppose $\#(2^{\Omega_{n-1}}) = 2_{n-1}$. Observe that $2^{\Omega_n} = \{\{x_n\} \cup A, A \in 2^{\Omega_{n-1}}\} \cup 2^{\Omega_{n-1}}\}$ hence $\#(2^{\Omega_n}) = 2\#(2^{\Omega_{n-1}}) = 2^n$. This proves finiteness. To show that 2^Ω is a σ-algebra we check:

1. $\emptyset \subset \Omega$ hence $\emptyset \in 2^\Omega$.
2. If $A \in 2^\Omega$ then $A \subset \Omega$ and $A^c \subset \Omega$ hence $A^c \in 2^\Omega$.
3. Let $(A_n)_{n\geq 1}$ be a sequence of subsets of Ω. Then $\bigcup_{n=1}^\infty A_n$ is also a subset of Ω hence in 2^Ω.

Therefore 2^Ω is a σ-algebra.

2.2 We check if $\mathcal{H} = \cap_{\alpha\in A}\mathcal{G}_\alpha$ has the three properties of a σ-algebra:

1. $\emptyset \in \mathcal{G}_\alpha$ $\forall\alpha \in A$ hence $\emptyset \in \cap_{\alpha\in A}\mathcal{G}_\alpha$.
2. If $B \in \cap_{\alpha\in A}\mathcal{G}_\alpha$ then $B \in \mathcal{G}_\alpha$ $\forall\alpha \in A$. This implies that $B^c \in \mathcal{G}_\alpha$ $\forall\alpha \in A$ since each $\mathcal{G}_\alpha$ is a σ-algebra. So $B^c \in \cap_{\alpha\in A}\mathcal{G}_\alpha$.
3. Let $(A_n)_{n\geq 1}$ be a sequence in $\mathcal{H}$. Since each $A_n \in \mathcal{G}_\alpha$, $\bigcup_{n=1}^\infty A_n$ is in $\mathcal{G}_\alpha$ since $\mathcal{G}_\alpha$ is a σ-algebra for each $\alpha \in A$. Hence $\bigcup_{n=1}^\infty A_n \in \cap_{\alpha\in A}\mathcal{G}_\alpha$.

Therefore $\mathcal{H} = \cap_{\alpha\in A}\mathcal{G}_\alpha$ is a σ-algebra.

2.3 a. Let $x \in (\cup_{n=1}^\infty A_n)^c$. Then $x \in A_n^c$ for all n, hence $x \in \cap_{n=1}^\infty A_n^c$. So $(\cup_{n=1}^\infty A_n)^c \subset \cap_{n=1}^\infty A_n^c$. Similarly if $x \in \cap_{n=1}^\infty A_n^c$ then $x \in A_n^c$ for any n hence $x \in (\cup_{n=1}^\infty A_n)^c$. So $(\cup_{n=1}^\infty A_n)^c = \cap_{n=1}^\infty A_n^c$.

b. By part-a $\cap_{n=1}^\infty A_n = (\cup_{n=1}^\infty A_n^c)^c$, hence $(\cap_{n=1}^\infty A_n)^c = \cup_{n=1}^\infty A_n^c$.

2.4 $\liminf_{n\to\infty} A_n = \cup_{n=1}^\infty B_n$ where $B_n = \cap_{m\geq n}A_m \in \mathcal{A}$ $\forall n$ since $\mathcal{A}$ is closed under taking countable intersections. Therefore $\liminf_{n\to\infty} A_n \in \mathcal{A}$ since $\mathcal{A}$ is closed under taking countable unions.

By De Morgan's Law it is easy to see that $\limsup A_n = (\liminf_{n\to\infty} A_n^c)^c$, hence $\limsup_{n\to\infty} A_n \in \mathcal{A}$ since $\liminf_{n\to\infty} A_n^c \in \mathcal{A}$ and $\mathcal{A}$ is closed under taking complements.

Note that $x \in \liminf_{n\to\infty} A_n \Rightarrow \exists n^*$ s.t $x \in \cap_{m\geq n^*}A_m \Rightarrow x \in \cap_{m\geq n}A_m \forall n \Rightarrow x \in \limsup_{n\to\infty} A_n$. Therefore $\liminf_{n\to\infty} A_n \subset \limsup_{n\to\infty} A_n$.

2.8 Let $\mathcal{L} = \{B \subset \mathbf{R} : f^{-1}(B) \in \mathcal{B}\}$. It is easy to check that $\mathcal{L}$ is a σ-algebra. Since f is continuous $f^{-1}(B)$ is open (hence Borel) if B is open. Therefore $\mathcal{L}$ contains the open sets which implies $\mathcal{L} \supset \mathcal{B}$ since $\mathcal{B}$ is generated by the open sets of $\mathbf{R}$. This proves that $f^{-1}(B) \in \mathcal{B}$ if $B \in \mathcal{B}$ and that $\mathcal{A} = \{A \subset \mathbf{R} : \exists B \in \mathcal{B}$ with $A = f^{-1}(B) \in \mathcal{B}\} \subset \mathcal{B}$.

Solutions to selected problems of Chapter 3

3.7 a. Since $P(B) > 0$ $P(.|B)$ defines a probability measure on $\mathcal{A}$, therefore by Theorem 2.4 $\lim_{n\to\infty} P(A_n|B) = P(A|B)$.

b. We have that $A \cap B_n \to A \cap B$ since $\mathbf{1}_{A\cap B_n}(w) = \mathbf{1}_A(w)\mathbf{1}_{B_n}(w) \to \mathbf{1}_A(w)\mathbf{1}_B(w)$. Hence $P(A \cap B_n) \to P(A \cap B)$. Also $P(B_n) \to P(B)$. Hence

$$P(A|B_n) = \frac{P(A \cap B_n)}{P(B_n)} \to \frac{P(A \cap B)}{P(B)} = P(A|B).$$

c.

$$P(A_n|B_n) = \frac{P(A_n \cap B_n)}{P(B_n)} \to \frac{P(A \cap B)}{P(B)} = P(A|B)$$

since $A_n \cap B_n \to A \cap B$ and $B_n \to B$.

3.11 Let $B = \{x_1, x_2, \ldots, x_b\}$ and $R = \{y_1, y_2, \ldots, y_r\}$ be the sets of b blue balls and r red balls respectively. Let $B' = \{x_{b+1}, x_{b+2}, \ldots, x_{b+d}\}$ and $R' = \{y_{r+1}, y_{r+2}, \ldots, y_{r+d}\}$ be the sets of d-new blue balls and d-new red balls respectively. Then we can write down the sample space Ω as

$$\Omega = \{(a, b) : (a \in B \text{ and } b \in B \cup B' \cup R) \text{ or } (a \in R \text{ and } b \in R \cup R' \cup B)\}.$$

Clearly $\text{card}(\Omega) = b(b + d + r) + r(b + d + r) = (b + r)(b + d + r)$. Now we can define a probability measure P on 2^Ω by

$$P(A) = \frac{\text{card}(A)}{\text{card}(\Omega)}.$$

a. Let

$$\begin{aligned} A &= \{ \text{ second ball drawn is blue}\} \\ &= \{(a, b) : a \in B, b \in B \cup B'\} \cup \{(a, b) : a \in R, b \in B\} \end{aligned}$$

$\text{card}(A) = b(b + d) + rb = b(b + d + r)$, hence $P(A) = \frac{b}{b+r}$.

b. Let

$$\begin{aligned} B &= \{ \text{ first ball drawn is blue}\} \\ &= \{(a, b) \in \Omega : a \in B\} \end{aligned}$$

Observe $A \cap B = \{(a, b) : a \in B, b \in B \cup B'\}$ and $\text{card}(A \cap B) = b(b + d)$. Hence

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{\text{card}(A \cap B)}{\text{card}(A)} = \frac{b + d}{b + d + r}.$$

3.17 We will use the inequality $1 - x > e^{-x}$ for $x > 0$, which is obtained by taking Taylor's expansion of e^{-x} around 0.

$$\begin{aligned} P((A_1 \cup \ldots \cup A_n)^c) &= P(A_1^c \cap \ldots \cap A_n^c) \\ &= (1 - P(A_1)) \ldots (1 - P(A_n)) \\ &\leq \exp(-P(A_1)) \ldots \exp(-P(A_n)) = \exp(-\sum_{i=1}^{n} P(A_i)) \end{aligned}$$

Solutions to selected problems of Chapter 4

4.1 Observe that

$$\begin{aligned} P(\text{k successes}) &= \begin{pmatrix} n \\ 2 \end{pmatrix} \frac{\lambda}{n}^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \\ &= C a_n b_{1,n} \dots b_{k,n} d_n \end{aligned}$$

where

$$C = \frac{\lambda^k}{k!} \quad a_n = (1 - \frac{\lambda}{n})^n \quad b_{j,n} = \frac{n - j + 1}{n} \quad d_n = (1 - \frac{\lambda}{n})^{-k}$$

It is clear that $b_{j,n} \to 1 \; \forall j$ and $d_n \to 1$ as $n \to \infty$. Observe that

$$\log((1 - \frac{\lambda}{n})^n) = n(\frac{\lambda}{n} - \frac{\lambda^2}{n^2}\frac{1}{\xi^2}) \text{ for some } \xi \in (1 - \tfrac{\lambda}{n}, 1)$$

by Taylor series expansion of $\log(x)$ around 1. It follows that $a_n \to e^{-\lambda}$ as $n \to \infty$ and that

$$|\text{Error}| = |e^{n \log(1 - \frac{\lambda}{n})} - e^{-\lambda}| \geq |n \log(1 - \frac{\lambda}{n}) - \lambda| = n \frac{\lambda^2}{n^2} \frac{1}{\xi^2} \geq \lambda p$$

Hence in order to have a good approximation we need n large and p small as well as λ to be of moderate size.

5.7 We put $x_n = P(X \text{ is even})$ for $X \sim B(p, n)$. Let us prove by induction that $x_n = \frac{1}{2}(1 + (1-2p)^n)$. For $n = 1$, $x_1 = 1 - p = \frac{1}{2}(1 + (1-2p)^1)$. Assume the formula is true for $n - 1$. If we condition on the outcome of the first trial we can write

$$\begin{aligned} x_n &= p(1 - x_{n-1}) + (1-p)x_n \\ &= p(1 - \frac{1}{2}(1 + (1-2p)^{n-1})) + (1-p)(\frac{1}{2}(1 + (1-2p)^{n-1})) \\ &= \frac{1}{2}(1 + (1-2p)^n) \end{aligned}$$

hence we have the result.

5.11 Observe that $E(|X - \lambda|) = \sum_{i<\lambda}(\lambda - i)p_i + \sum_{i\geq\lambda}(i-\lambda)p_i$. Since $\sum_{i\geq\lambda}(i-\lambda)p_i = \sum_{i=0}^{\infty}(i-\lambda)p_i - \sum_{i<\lambda}(i-\lambda)p_i$ we have that $E(|X-\lambda|) = 2\sum_{i<\lambda}(\lambda - i)p_i$. So

$$\begin{aligned} E(|X-\lambda|) &= 2\sum_{i<\lambda}(\lambda - i)p_i \\ &= 2\sum_{i=1}^{\lambda-1}(\lambda - i)\frac{e^{-\lambda}\lambda^k}{k!} \\ &= 2e^{-\lambda}\sum_{i=0}^{\lambda-1}(\frac{\lambda^{k+1}}{k!} - \frac{\lambda^k}{(k-1)!}) \\ &= 2e^{-\lambda}\frac{\lambda^\lambda}{(k-1)!}. \end{aligned}$$

7.1 Suppose $\lim_{n\to\infty} P(A_n) \neq 0$. Then there exists $\epsilon > 0$ such that there are distinct $A_{n_1}, A_{n_2}, \dots$ with $P(A_{n_k}) > 0$ for every $k \leq 1$. This gives $\sum_{k=1}^{\infty} P(A_{n_k}) = \infty$ which is a contradiction since by the hypothesis that the A_n are disjoint we have that $\sum_{k=1}^{\infty} P(A_{n_k}) = P(\cup_{n=1}^{\infty} A_{n_k}) \leq 1$.

7.2 Let $\mathcal{A}_n = \{A_\beta : P(A_\beta) > 1/n\}$. $\mathcal{A}_n$ is a finite set otherwise we can pick disjoint $A_{\beta_1}, A_{\beta_2}, \dots$ in $\mathcal{A}_n$. This would give us $P \cup_{m=1}^{\infty} A_{\beta_m} = \sum_{m=1}^{\infty} P(A_{\beta_m}) = \infty$ which is a contradiction. Now $\{A_\beta : \beta \in B\} = \cup_{n=1}^{\infty} \mathcal{A}_n$ hence $(A_\beta)_{\beta \in B}$ is countable since it is a countable union of finite sets.

7.11 Note that $\{x_0\} = \cap_{n=1}^{\infty} [x_0 - 1/n, x_0]$ therefore $\{x_0\}$ is a Borel set. $P(\{x_0\}) = \lim_{n\to\infty} P([x_0 - 1/n, x_0])$. Assuming that f is continuous we have that f is bounded by some M on the interval $[x_0 - 1/n, x_0]$ hence $P(\{x_0\}) = \lim_{n\to\infty} M(1/n) = 0$.
Remark: In order this result to be true we don't need f to be continuous. When we define the Lebesgue integral (or more generally integral with respect to a measure) and study its properties we will see that this result is true for all Borel measurable non-negative f.

7.16 First observe that $F(x) - F(x-) > 0$ iff $P(\{x\}) > 0$. The family of events $\{\{x\} : P(\{x\}) > 0\}$ can be at most countable as we have proven in problem 7.2 since these events are disjoint and have positive probability. Hence F can have at most countable discontinuities. For an example with infinitely many jump discontinuities consider the Poisson distribution.

7.18 Let F be as given. It is clear that F is a nondecreasing function. For $x < 0$ and $x \geq 1$ right continuity of F is clear. For any $0 < x < 1$ let i^* be such that $\frac{1}{i^*+1} \leq x < \frac{1}{i*}$. If $x_n \downarrow x$ then there exists N such that $\frac{1}{i^*+1} \leq x_n < \frac{1}{i*}$ for every $n \geq N$. Hence $F(x_n) = F(x)$ for every $n \geq N$ which implies that F is right continuous at x. For $x = 0$ we have that $F(0) = 0$. Note that for any ϵ there exists N such that $\sum_{i=N}^{\infty} \frac{1}{2^i} < \epsilon$. So for all x s.t. $|x| \leq \frac{1}{N}$ we have that $F(x) \leq \epsilon$. Hence $F(0+) = 0$. This proves the right continuity of F for all x. We also have that $F(\infty) = \sum_{i=1}^{\infty} \frac{1}{2^i} = 1$ and $F(-\infty) = 0$ so F is a distribution function of a probability on $\mathbf{R}$.

a. $P([1, \infty)) = F(\infty) - F(1-) = 1 - \sum_{n=2}^{\infty} = 1 - \frac{1}{2} = \frac{1}{2}$.
b. $P([\frac{1}{10}, \infty)) = F(\infty) - F(\frac{1}{10}-) = 1 - \sum_{n=11}^{\infty} \frac{1}{2^i} = 1 - 2^{-10}$.
c $P(\{0\}) = F(0) - F(0-) = 0$.
d. $P([0, \frac{1}{2})) = F(\frac{1}{2}-) - F(0-) = \sum_{n=3}^{\infty} \frac{1}{2^i} - 0 = \frac{1}{4}$.
e. $P((-\infty, 0)) = F(0-) = 0$.
f. $P((0, \infty)) = 1 - F(0) = 1$.

Solutions to selected problems of Chapter 9

9.1 It is clear by the definition of $\mathcal{F}$ that $X^{-1}(B) \in \mathcal{F}$ for every $B \in \mathcal{B}$. So X is measurable from $(\Omega, \mathcal{F})$ to $(\mathbf{R}, \mathcal{B})$.

9.2 Since X is both $\mathcal{F}$ and $\mathcal{G}$ measurable for any $B \in \mathcal{B}$, $P(X \in B) = P(X \in B)P(X \in B) = 0$ or 1. Without loss of generality we can assume that there exists a closed interval I such that $P(I) = 1$. Let $\Lambda_n = \{t_0^n, \dots t_{l_n}^n\}$ be a partition of I such that $\Lambda_n \subset \Lambda_{n+1}$ and $\sup_k t_k^n - t_{k-1}^n \to 0$. For each n there exists $k^*(n)$ such that $P(X \in [t_{k^*}^n, t_{k^*+1}^n]) = 1$ and $[t_{k^*(n+1)}^n, t_{k^*(n+1)+1}^n] \subset [t_{k^*(n)}^n, t_{k^*(n)+1}^n]$. Now $a_n = t_{k^*(n)}^n$ and $b_n = t_{k^*(n)}^n + 1$ are both Cauchy sequences with a common limit c. So $1 = \lim_{n\to\infty} P(X \in (t_{k^*}^n, t_{k^*+1}^n]) = P(X = c)$.

9.3 $X^{-1}(A) = (Y^{-1}(A) \cap (Y^{-1}(A) \cap X^{-1}(A)^c)^c) \cup (X^{-1}(A) \cap Y^{-1}(A)^c)$. Observe that both $Y^{-1}(A) \cap (X^{-1}(A))^c$ and $X^{-1}(A) \cap Y^{-1}(A)^c$ are null sets and therefore measurable. Hence if $Y^{-1}(A) \in \mathcal{A}'$ then $X^{-1}(A) \in \mathcal{A}'$. In other words if Y is $\mathcal{A}'$ measurable so is X.

9.4 Since X is integrable, for any $\epsilon > 0$ there exists M such that $\int |X| \mathbf{1}_{\{X>M\}} dP < \epsilon$ by the dominated convergence theorem. Note that

$$\begin{aligned} E[X\mathbf{1}_{A_n}] &= E[X\mathbf{1}_{A_n}\mathbf{1}_{\{X>M\}}] + E[X\mathbf{1}_{A_n}\mathbf{1}_{\{X\le M\}}] \\ &\le E[|X|\mathbf{1}_{\{X\le M\}}] + MP(A_n) \end{aligned}$$

Since $P(A_n) \to 0$, there exists N such that $P(A_n) \le \frac{\epsilon}{M}$ for every $n \ge N$. Therefore $E[X\mathbf{1}_{A_n}] \le \epsilon + \epsilon \ \forall n \ge N$, i.e. $\lim_{n\to\infty} E[X\mathbf{1}_{A_n}] = 0$.

9.5 It is clear that $0 \le Q(A) \le 1$ and $Q(\Omega) = 1$ since X is nonnegative and $E[X] = 1$. Let $A_1, A_2, \dots$ be disjoint. Then

$$Q(\cup_{n=1}^\infty A_n) = E[X\mathbf{1}_{\cup_{n=1}^\infty A_n}] = E[\sum_{n=1} X\mathbf{1}_{A_n}] = \sum_{n=1}^\infty E[X\mathbf{1}_{A_n}]$$

where the last equality follows from the monotone convergence theorem. Hence $Q(\cup_{n=1}^\infty A_n) = \sum_{n=1}^\infty Q(A_n)$. Therefore Q is a probability measure.

9.6 If $P(A) = 0$ then $X\mathbf{1}_A = 0$ a.s. Hence $Q(A) = E[X\mathbf{1}_A] = 0$. Now assume P is the uniform distribution on $[0,1]$. Let $X(x) = 2\mathbf{1}_{[0,1/2]}(x)$. Corresponding measure Q assigns zero measure to $(1/2, 1]$, however $P((1/2, 1]) = 1/2 \neq 0$.

9.7 Let's prove this first for simple functions, i.e. let Y be of the form

$$Y = \sum_{i=1}^n c_i \mathbf{1}_{A_i}$$

for disjoint $A_1, \ldots, A_n$. Then

$$E_Q[Y] = \sum_{i=1}^{n} c_i Q(A_i) = \sum_{i=1}^{n} c_i E[X\mathbf{1}_{A_i}] = E_P[XY]$$

For non-negative Y we take a sequence of simple functions $Y_n \uparrow Y$. Then

$$E_Q[Y] = \lim_{n\to\infty} E_Q[Y_n] = \lim_{n\to\infty} E_P[XY_n] = E_P[XY]$$

where the last equality follows from the monotone convergence theorem. For general $Y \in Ł^1(Q)$ we have that $E_Q[Y] = E_Q[Y^+] - E_Q[Y^-] = E_P[(XY)^+] - E_Q[(XY)^-] = E_P[XY]$.

9.8 a. Note that $\frac{1}{X}X = 1$ a.s. since $P(X > 0) = 1$. By problem 9.7 $E_Q[\frac{1}{X}] = E_P[\frac{1}{X}X] = 1$. So $\frac{1}{X}$ is Q-integrable.

b. $R : \mathcal{A} \to \mathbf{R}$, $R(A) = E_Q[\frac{1}{X}\mathbf{1}_A]$ is a probability measure since $\frac{1}{X}$ is non-negative and $E_Q[\frac{1}{X}] = 1$. Also $R(A) = E_Q[\frac{1}{X}\mathbf{1}_A] = E_P[\frac{1}{X}X\mathbf{1}_A] = P(A)$. So $R = P$.

9.9 Since $P(A) = E_Q[\frac{1}{X}\mathbf{1}_A]$ we have that $Q(A) = 0 \Rightarrow P(A) = 0$. Now combining the results of the previous problems we can easily observe that $Q(A) = 0 \Leftrightarrow P(A) = 0$ iff $P(X > 0) = 1$.

9.17. Let

$$g(x) = \frac{((x-\mu)b + \sigma)^2}{\sigma^2(1+b^2)^2}.$$

Observe that $\{X \geq \mu + b\sigma\} \in \{g(X) \geq 1\}$. So

$$P(\{X \geq \mu + b\sigma\}) \leq P(\{g(X) \geq 1\}) \leq \frac{E[g(X)]}{1}$$

where the last inequality follows from Markov's inequality. Since $E[g(X)] = \frac{\sigma^2(1+b^2)}{\sigma^2(1+b^2)^2}$ we get that

$$P(\{X \geq \mu + b\sigma\}) \leq \frac{1}{1+b^2}.$$

9.19

$$\begin{aligned} xP(\{X > x\}) &\leq E[X\mathbf{1}_{\{}X > x\}] \\ &= \int_x^\infty \frac{z}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \\ &= \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} \end{aligned}$$

Hence

$$P(\{X > x\}) \leq \frac{e^{-\frac{x^2}{2}}}{x\sqrt{2\pi}}$$

.

9.21 $h(t+s) = P(\{X > t+s\}) = P(\{X > t+s, X > s\}) = P(\{X > t+s | X > s\})P(\{X > s\}) = h(t)h(s)$ for all $t, s > 0$. Note that this gives $h(\frac{1}{n}) = h(1)^{\frac{1}{n}}$ and $h(\frac{m}{n}) = h(1)^{\frac{m}{n}}$. So for all rational r we have that $h(r) = \exp(\log(h(1))r)$. Since h is right continuous this gives $h(x) = \exp(\log(h(1))x)$ for all $x > 0$. Hence X has exponential distribution with parameter $-\log h(1)$.

Solutions to selected problems of Chapter 10

10.5 Let P be the uniform distribution on $[-1/2, 1/2]$. Let $X(x) = \mathbf{1}_{[-1/4,1/4]}$ and $Y(x) = \mathbf{1}_{[-1/4,1/4]^c}$. It is clear that $XY = 0$ hence $E[XY] = 0$. It is also true that $E[X] = 0$. So $E[XY] = E[X]E[Y]$ however it is clear that X and Y are not independent.

10.6 a. $P(\min(X,Y) > i) = P(X > i)P(Y > i) = \frac{1}{2^i}\frac{1}{2^i} = \frac{1}{4^i}$. So $P(\min(X,Y) \leq i) = 1 - P(\min(X,Y) > i) = 1 - \frac{1}{4^i}$.

b. $P(X = Y) = \sum_{i=1}^{\infty} P(X = i)P(Y = i) = \sum_{i=1}^{\infty} \frac{1}{2^i}\frac{1}{2^i} = \frac{1}{1-\frac{1}{4^i}} - 1 = \frac{1}{3}$.

c. $P(Y > X) = \sum_{i=1}^{\infty} P(Y > i)P(X = i) = \sum_{i=1}^{\infty} \frac{1}{2^i}\frac{1}{2^i} = \frac{1}{3}$.

d. $P(X \text{ divides } Y) = \sum_{i=1}^{\infty}\sum_{k=1}^{\infty} \frac{1}{2^i}\frac{1}{2^{ki}} = \sum_{i=1}^{\infty} \frac{1}{2^i}\frac{1}{2^i-1}$.

e. $P(X \geq kY) = \sum_{i=1}^{\infty} P(X \geq ki)P(Y = i) = \sum_{i=1}^{\infty} \frac{1}{2^i}\frac{1}{2^{ki}-1} = \frac{2}{2^{k+1}-1}$.

11.11. Since $P\{X > 0\} = 1$ we have that $P\{Y < 1\} = 1$. So $F_Y(y) = 1$ for $y \geq 1$. Also $P\{Y \leq 0\} = 0$ hence $F_Y(y) = 0$ for $y \leq 0$. For $0 < y < 1$ $P\{Y > y\} = P\{X < \frac{1-y}{y}\} = F_X(\frac{1-y}{y})$. So

$$F_Y(y) = 1 - \int_0^{\frac{1-y}{y}} f_X(x)dx \quad = 1- \quad \int_0^y \frac{-1}{z^2} f_X(\frac{1-z}{z})dz$$

by change of variables. Hence

$$f_Y(y) \quad = \quad \begin{cases} 0 & -\infty < y \leq 0 \\ \frac{1}{y^2} f_X(\frac{1-y}{y}) & 0 < y \leq 1 \\ 0 & 1 \leq y < \infty \end{cases}$$

11.15 Let $G(u) = \inf\{x : F(x) \geq u\}$. We would like to show $\{u : G(u) > y\} = \{u : F(Y) < u\}$. Let u be such that $G(u) > y$. Then $F(y) < u$ by definition of G. Hence $\{u : G(u) > y\} \subset \{u : F(Y) < u\}$. Now let u be such that $F(y) < u$. Then $y < x$ for any x such that $F(x) \geq u$ by monotonicity of F. Now by right continuity and the monotonicity of F we have that $F(G(u)) = \inf_{F(x) \geq u} F(x) \geq u$. Then by the previous statement $y < G(u)$. So $\{u : G(u) > y\} = \{u : F(Y) < u\}$. Now $P\{G(U) > y\} = P\{U > F(y)\} = 1 - F(y)$ so $G(U)$ has the desired distribution. **Remark:We only assumed the right continuity of F.**

12.6 Let $Z = (\frac{1}{\sigma_Y})Y - (\frac{\rho_{XY}}{\sigma_X})X$. Then $\sigma_Z^2 = (\frac{1}{\sigma_Y^2})\sigma_Y^2 - (\frac{\rho_{XY}^2}{\sigma_X^2})\sigma_X^2 - 2(\frac{\rho_{XY}}{\sigma_X\sigma_Y})\mathrm{Cov}(X,Y) = 1 - \rho_{XY}^2$. Note that $\rho_{XY} = \mp 1$ implies $\sigma_Z^2 = 0$ which implies $Z = c$ a.s. for some constant c. In this case $X = \frac{\sigma_X}{\sigma_Y \rho_{XY}}(Y - c)$ hence X is an affine function of Y.

12.11 Consider the mapping $g(x,y) = (\sqrt{x^2+y^2}, \arctan(\frac{x}{y}))$. Let $S_0 = \{(x,y) : y = 0\}$, $S_1 = \{(x,y) : y > 0\}$, $S_2 = \{(x,y) : y < 0\}$. Note that $\cup_{i=0}^2 S_i = \mathbf{R^2}$ and $m_2(S_0) = 0$. Also for $i = 1,2$ $g : S_i \to \mathbf{R^2}$ is injective and continuously differentiable. Corresponding inverses are given by $g_1^{-1}(z,w) = (z\sin w, z\cos w)$ and $g_2^{-1}(z,w) = (z\sin w, -z\cos w)$. In both cases we have that $|J_{g_i^{-1}}(z,w)| = z$ hence by Corollary 12.1 the density of (Z,W) is given by

$$
\begin{aligned}
f_{Z,W}(z,w) &= (\frac{1}{2\pi\sigma^2}e^{\frac{-z^2}{2\sigma}}z + \frac{1}{2\pi\sigma^2}e^{\frac{-z^2}{2\sigma}}z)\mathbf{1}_{(-\frac{\pi}{2},\frac{\pi}{2})}(w)\mathbf{1}_{(0,\infty)}(z) \\
&= \frac{1}{\pi}\mathbf{1}_{(-\frac{\pi}{2},\frac{\pi}{2})}(w) * \frac{z}{\sigma^2}e^{\frac{-z^2}{2\sigma}}\mathbf{1}_{(0,\infty)}(z)
\end{aligned}
$$

as desired.

12.12 Let $\mathcal{P}$ be the set of all permutations of $\{1,\ldots,n\}$. For any $\pi \in \mathcal{P}$ let X^π be the corresponding permutation of X, i.e. $X_k^\pi = X_{\pi_k}$. Observe that

$$P(X_1^\pi \le x_1, \ldots, X_n^\pi \le x_n) = F(x_1)\ldots F(X_n)$$

hence the law of X^π and X coincide on a πsystem generating $\mathcal{B}^n$ therefore they are equal. Now let $\Omega_0 = \{(x_1,\ldots,x_n) \in \mathbf{R}^n : x_1 < x_2 < \ldots < x_n\}$. Since X_i are i.i.d and have continuous distribution $P_X(\Omega_0) = 1$. Observe that

$$P\{Y_1 \le y_1, \ldots, Y_n \le y_n\} = P(\cup_{\pi\in\mathcal{P}}\{X_1^\pi \le y_1, \ldots, X_n^\pi \le y_n\} \cap \Omega_0)$$

Note that $\{X_1^\pi \le y_1, \ldots, X_n^\pi \le y_n\} \cap \Omega_0$, $\pi \in \mathcal{P}$ are disjoint and $P(\Omega_0 = 1)$ hence

$$
\begin{aligned}
P\{Y_1 \le y_1, \ldots, Y_n \le y_n\} &= \sum_{\pi\in\mathcal{P}} P\{X_1^\pi \le y_1, \ldots, X_n^\pi \le y_n\} \\
&= n!F(y_1)\ldots F(y_n)
\end{aligned}
$$

for $y_1 \le \ldots \le y_n$. Hence

$$
f_Y(y_1,\ldots,y_n) = \begin{cases} n!f(y_1)\ldots f(y_n) & y_1 \le \ldots \le y_n \\ 0 & \text{otherwise} \end{cases}
$$

14.7 $\varphi_X(u)$ is real valued iff $\varphi_X(u) = \overline{\varphi}_X(u) = \varphi_{-X}(u)$. By uniqueness theorem $\varphi_X(u) = \varphi_{-X}(u)$ iff $F_X = F_{-X}$. Hence $\varphi_X(u)$ is real valued iff $F_X = F_{-X}$.

14.9 We use induction. It is clear that the statement is true for $n = 1$. Put $Y_n = \sum_{i=1}^n X_i$ and assume that $E[(Y_n)^3] = \sum_{i=1}^n E[(X_i)^3]$. Note that this implies $\frac{d^3}{dx^3}\varphi_{Y_n}(0) = -i\sum_{i=1}^n E[(X_i)^3]$. Now $E[(Y_{n+1})^3] = E[(X_{n+1}+Y_n)^3] = -i\frac{d^3}{dx^3}(\varphi_{X_{n+1}}\varphi_{Y_n})(0)$ by independence of X_{n+1} and Y_n. Note that

$$\begin{aligned}
\frac{d^3}{dx^3}\varphi_{X_{n+1}}\varphi_{Y_n}(0) &= \frac{d^3}{dx^3}\varphi_{X_{n+1}}(0)\varphi_{Y_n}(0) \\
&\quad + 3\frac{d^2}{dx^2}\varphi_{X_{n+1}}(0)\frac{d}{dx}\varphi_{Y_n}(0) + 3\frac{d}{dx}\varphi_{X_{n+1}}(0)\frac{d^2}{dx^2}\varphi_{Y_n}(0) \\
&\quad + \varphi_{X_{n+1}}(0)\frac{d^3}{dx^3}\varphi_{Y_n}(0) \\
&= \frac{d^3}{dx^3}\varphi_{X_{n+1}}(0) + \frac{d^3}{dx^3}\varphi_{Y_n}(0) \\
&= -i\left(E[(X_{n+1})^3] + \sum_{i=1}^n E[(X_i)^3]\right)
\end{aligned}$$

where we used the fact that $\frac{d}{dx}\varphi_{X_{n+1}}(0) = iE(X_{n+1}) = 0$ and $\frac{d}{dx}\varphi_{Y_n}(0) = iE(Y_n) = 0$. So $E[(Y_{n+1})^3] = \sum_{i=1}^{n+1} E[(X_i)^3]$ hence the induction is complete.

14.10 It is clear that $0 \le \nu(A) \le 1$ since

$$0 \le \sum_{j=1}^n \lambda_j\mu_j(A) \le \sum_{j=1}^n \lambda_j = 1.$$

Also for A_i disjoint

$$\begin{aligned}
\nu(\cup_{i=1}^\infty A_i) &= \sum_{j=1}^n \lambda_j\mu_j(\cup_{i=1}^\infty A_i) \\
&= \sum_{j=1}^n \lambda_j \sum_{i=1}^\infty \mu_j(A_i) \\
&= \sum_{i=1}^\infty \sum_{j=1}^n \lambda_j\mu_j(A_i) \\
&= \sum_{i=1}^\infty \nu(A_i)
\end{aligned}$$

Hence ν is countably additive therefore it is a probability mesure. Note that $\int \mathbf{1}_A d\nu(dx) = \sum_{j=1}^{n} \lambda_j \int \mathbf{1}_A(x) d\mu_j(dx)$ by definition of ν. Now by linearity and monotone convergence theorem for a non-negative Borel function f we have that $\int f(x)\nu(dx) = \sum_{j=1}^{n} \lambda_j \int f(x) d\mu_j(dx)$. Extending this to integrable f we have that $\hat{\nu}(u) = \int e^{iux}\nu(dx) = \sum_{j=1}^{n} \lambda_j \int e^{iux} d\mu_j(dx) = \sum_{j=1}^{n} \lambda_j \hat{\mu_j}(u)$.

14.11 Let ν be the double exponential distribution, μ_1 be the distribution of Y and μ_2 be the distribution of $-Y$ where Y is an exponential r.v. with parameter $\lambda = 1$. Then we have that $\nu(A) = \frac{1}{2}\int_{A\cap(0,\infty)} e^{-x}dx + \frac{1}{2}\int_{A\cap(-\infty,0)} e^{x}dx = \frac{1}{2}\mu_1(A) + \frac{1}{2}\mu_2(A)$. By the previous exercise we have that $\hat{\nu}(u) = \frac{1}{2}\hat{\mu}_1(u) + \frac{1}{2}\hat{\mu}_2(u) = \frac{1}{2}(\frac{1}{1-iu} + \frac{1}{1+iu}) = \frac{1}{1+u^2}$.

14.15. Note that $E\{X^n\} = (-i)^n \frac{d^n}{dx^n}\varphi_X(0)$. Since $X \sim N(0,1)$ $\varphi_X(s) = e^{-s^2/2}$. Note that we can get the derivatives of any order of $e^{-s^2/2}$ at 0 simply by taking Taylor's expansion of e^x:

$$\begin{aligned} e^{-s^2/2} &= \sum_{i=0}^{\infty} \frac{(-s^2/2)^n}{n!} \\ &= \sum_{i=0}^{\infty} \frac{1}{2n!} \frac{(-i)^{2n}(2n)!}{2^n n!} s^{2n} \end{aligned}$$

hence $E\{X^n\} = (-i)^n \frac{d^n}{dx^n}\varphi_X(0) = 0$ for n odd. For $n = 2k$ $E\{X^{2k}\} = (-i)^{2k} \frac{d^{2k}}{dx^{2k}}\varphi_X(0) = (-i)^{2k}\frac{(-i)^{2k}(2k)!}{2^k k!} = \frac{(2k)!}{2^k k!}$ as desired.

Solutions to selected problems of Chapter 15

15.1 a. $E\{\overline{x}\} = \frac{1}{n}\sum_{i=1}^{n} E\{X_i\} = \mu$.

b. Since $X_1, \ldots, X_n$ are independent $\text{Var}(\overline{x}) = \frac{1}{n^2}\sum_{i=1}^{n} \text{Var}\{X_i\} = \frac{\sigma^2}{n}$.

c. Note that $S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i)^2 - \overline{x}^2$. Hence $E(S^2) = \frac{1}{n}\sum_{i=1}^{n}(\sigma^2 + \mu^2) - (\frac{\sigma^2}{n} + \mu^2) = \frac{n-1}{n}\sigma^2$.

15.17 Note that $\varphi_Y(u) = \prod_{i=1}^{\alpha} \varphi_{X_i}(u) = (\frac{\beta}{\beta - iu})^{\alpha}$ which is the characteristic function of Gamma(α,β) random variable. Hence by uniqueness of characteristic function Y is Gamma(α,β).

Solutions to selected problems of Chapter 16

16.3 $P(\{Y \le y\}) = P(\{X \le y\} \cap \{Z = 1\}) + P(\{-X \le y\} \cap \{Z = -1\}) = \frac{1}{2}\Phi(y) + \frac{1}{2}\Phi(-y) = \Phi(y)$ since Z and X are independent and $\Phi(y)$ is symmetric. So Y is normal. Note that $P(X + Y = 0) = \frac{1}{2}$ hence $X + Y$ can not be normal. So (X, Y) is not Gaussian even though both X and Y are normal.

16.4 Observe that

$$Q = \sigma_X\sigma_Y \begin{bmatrix} \frac{\sigma_X}{\sigma_Y} & \rho \\ \rho & \frac{\sigma_Y}{\sigma_X} \end{bmatrix}$$

So $\det(Q) = \sigma_X\sigma_Y(1 - \rho^2)$. So $\det(Q) = 0$ iff $\rho = \mp 1$. By Corollary 16.2 the joint density of (X, Y) exists iff $-1 < \rho < 1$. (By Cauchy-Schwartz we know that $-1 \le \rho \le 1$). Note that

$$Q^{-1} = \frac{1}{\sigma_X\sigma_Y(1-\rho^2)} \begin{matrix} \frac{\sigma_Y}{\sigma_X} & -\rho \\ -\rho & \frac{\sigma_X}{\sigma_Y} \end{matrix}$$

Substituting this in formula 16.5 we get that

$$\begin{aligned} f_{(X,Y)}(x, y) &= \frac{1}{2\pi\sigma_X\sigma_Y(1-\rho^2)} \exp\left\{ \frac{-1}{2(1-\rho^2)} \left(\left(\frac{x - \mu_X}{\sigma_X} \right)^2 \right.\right. \\ &\quad \left.\left. - \frac{2\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} + \left(\frac{y-\mu_Y}{\sigma_Y} \right)^2 \right) \right\}. \end{aligned}$$

16.6 By Theorem 16.2 there exists a multivariate normal r.v. Y with $E(Y) = 0$ and a diagonal covariance matrix Λ s.t. $X - \mu = AY$ where A is an orthogonal matrix. Since $Q = A\Lambda A^*$ and $\det(Q) > 0$ the diagonal entries of Λ are strictly positive hence we can define $B = \Lambda^{-1/2}A^*$. Now the covariance matrix $\tilde{Q}$ of $B(X - \mu)$ is given by

$$\begin{aligned} \tilde{Q} &= \Lambda^{-1/2}A^*A\Lambda A^*A\Lambda^{-1/2} \\ &= I \end{aligned}$$

So $B(X - \mu)$ is standard normal.

16.17 We know that as in Exercise 16.6 if $B = \Lambda^{-1/2}A^*$ where A is the orthogonal matrix s.t. $Q = A\Lambda A^*$ then $B(X - \mu)$ is standard normal. Note that this gives $(X - \mu)^*Q^{-1}(X - \mu) = (X - \mu)^*B^*B(X - \mu)$ which has chi-square distribution with n degrees of freedom.

17.1 Let $n(m)$ and $j(m)$ be such that $Y_m = n(m)^{1/p} Z_{n(m),j(m)}$. This gives that $P(|Y_m| > 0) = \frac{1}{n(m)} \to 0$ as $m \to \infty$. So Y_m converges to 0 in probability. However $E[|Y_m|^p] = E[n(m) Z_{n(m),j(m)}] = 1$ for all m. So Y_m does not converge to 0 in L^p.

17.2 Let $X_n = 1/n$. It is clear that X_n converge to 0 in probability. If $f(x) = \mathbf{1}_{\{0\}}(x)$ then we have that $P(|f(X_n) - f(0)| > \epsilon) = 1$ for every $\epsilon \geq 1$, so $f(X_n)$ does not converge to $f(0)$ in probability.

17.3 First observe that $E(S_n) = \sum_{i=1}^n E(X_n) = 0$ and that $\text{Var}(S_n) = \sum_{i=1}^n \text{Var}(X_n) = n$ since $E(X_n) = 0$ and $\text{Var}(X_n) = E(X_n^2) = 1$. By Chebyshev's inequality $P(|\frac{S_n}{n}| \geq \epsilon) = P(|S_n| \geq n\epsilon) \leq \frac{\text{Var}(S_n)}{n^2\epsilon^2} = \frac{n}{n^2\epsilon^2} \to 0$ as $n \to \infty$. Hence $\frac{S_n}{n}$ converges to 0 in probability.

17.4 Note that Chebyshev's inequality gives $P(|\frac{S_{n^2}}{n^2}| \geq \epsilon) \leq \frac{1}{n^2\epsilon^2}$. Since $\sum_{i=1}^\infty \frac{1}{n^2\epsilon^2} < \infty$ by Borel Cantelli Theorem $P(\limsup_n \{|\frac{S_{n^2}}{n^2}| \geq \epsilon\}) = 0$. Let $\Omega_0 = \left(\cup_{m=1}^\infty \limsup_n \{|\frac{S_{n^2}}{n^2}| \geq \frac{1}{m}\}\right)^c$. Then $P(\Omega_0) = 1$. Now let's pick $w \in \Omega_0$. For any ϵ there exists m s.t. $\frac{1}{m} \leq \epsilon$ and $w \in (\limsup_n \{|\frac{S_{n^2}}{n^2}| \geq \frac{1}{m}\})^c$. Hence there are finitely many n s.t. $|\frac{S_{n^2}}{n^2}| \geq \frac{1}{m}$ which implies that there exists $N(w)$ s.t. $|\frac{S_{n^2}}{n^2}| \leq \frac{1}{m}$ for every $n \geq N(w)$. Hence $\frac{S_{n^2}(w)}{n^2} \to 0$. Since $P(\Omega_0) = 1$ we have almost sure convergence.

17.12 $Y < \infty$ a.s. which follows by Exercise 17.11 since $X_n < \infty$ and $X < \infty$ a.s. Let $Z = \frac{1}{c}\frac{1}{1+Y}$. Observe that $Z > 0$ a.s. and $E_P(Z) = 1$. Therefore as in Exercise 9.8 $Q(A) = E_P(Z\mathbf{1}_A)$ defines a probability measure and $E_Q(|X_n - X|) = E_P(Z|X_n - X|)$. Note that $Z|X_n - X| \leq 1$ a.s. and $X_n \to X$ a.s. by hypothesis, hence by dominated convergence theorem $E_Q(|X_n - X|) = E_P(Z|X_n - X|) \to 0$, i.e. X_n tends to X in L^1 with respect to Q.

17.14 First observe that $|E(X_n^2) - E(X^2)| \leq E(|X_n^2 - X^2|)$. Since $|X_n^2 - X^2| \leq (X_n - X)^2 + 2|X||X_n - X|$ we get that $|E(X_n^2) - E(X^2)| \leq E((X_n - X)^2) + 2E(|X||X_n - X|)$. Note that first term goes to 0 since X_n tends to X in L^2. Applying Cauchy Schwarz inequality to the second term we get $E(|X||X_n - X|) \leq \sqrt{E(X^2)E(|X_n - X|^2)}$, hence the second term also goes to 0 as $n \to \infty$. Now we can conclude $E(X_n^2) \to E(X^2)$.

17.15 For any $\epsilon > 0$ $P(\{|X| \leq c + \epsilon\}) \geq P(\{|X_n| \leq c, |X_n - X| \leq \epsilon\}) \to 1$ as $n \to \infty$. Hence $P(\{|X| \leq c + \epsilon\}) = 1$. Since $\{X \leq c\} = \cap_{m=1}^\infty \{X \leq c + \frac{1}{m}\}$ we get that $P\{X \leq c\} = 1$. Now we have that $E(|X_n - X|) = E(|X_n - X|\mathbf{1}_{\{|X_n - X| \leq \epsilon\}}) + E(|X_n - X|\mathbf{1}_{\{|X_n - X| > \epsilon\}}) \leq \epsilon + 2c(P\{|X_n - X| > \epsilon\})$, hence choosing n large we can make $E(|X_n - X|)$ arbitrarily small, so X_n tends to X in L^1.

18.8 Note that $\varphi_{Y_n}(u) = \Pi_{i=1}^n \varphi_{X_i}(\frac{u}{n}) = \Pi_{i=1}^n e^{-\frac{|u|}{n}} = e^{-|u|}$, hence Y_n is also Cauchy with $\alpha = 0$ and $\beta = 1$ which is independent of n, hence trivially Y_n converges in distribution to a Cauchy distributed r.v. with $\alpha = 0$ and $\beta = 1$. However Y_n does not converge to any r.v. in probability. To see this, suppose there exists Y s.t. $P(|Y_n - Y| > \epsilon) \to 0$. Note that $P(|Y_n - Y_m| > \epsilon) \leq P(|Y_n - Y| > \frac{\epsilon}{2}) + P(|Y_m - Y| > \frac{\epsilon}{2})$. If we let $m = 2n$, $|Y_n - Y_m| = \frac{1}{2}|\frac{1}{n}\sum_{i=1}^n X_i - \frac{1}{n}\sum_{i=n+1}^{2n} X_i|$ which is equal in distribution to $\frac{1}{2}|U - W|$ where U and W are independent Cauchy r.v.'s with $\alpha = 0$ and $\beta = 1$. Hence $P(|Y_n - Y_m| > \frac{\epsilon}{2})$ does not depend on n and does not converge to 0 if we let $m = 2n$ and $n \to \infty$ which is a contradiction since we assumed the right hand side converges to 0.

18.16 Define f_m as the following sequence of functions:

$$f_m(x) = \begin{cases} x^2 & \text{if } |x| \leq N - \frac{1}{m} \\ (N - \frac{1}{m})x - (N - \frac{1}{m})N & \text{if } x \geq N - \frac{1}{m} \\ -(N - \frac{1}{m})x + (N - \frac{1}{m})N & \text{if } x \leq -N + \frac{1}{m} \\ 0 & \text{otherwise} \end{cases}$$

Note that each f_m is continuous and bounded. Also $f_m(x) \uparrow \mathbf{1}_{(-N,N)}(x)x^2$ for every $x \in \mathbf{R}$. Hence

$$\int_{-N}^{N} x^2 F(dx) = \lim_{m\to\infty} \int_{-\infty}^{\infty} f_m(x)F(dx)$$

by monotone convergence theorem. Now

$$\int_{-\infty}^{\infty} f_m(x)F(dx) = \lim_{n\to\infty} \int_{-\infty}^{\infty} f_m(x)F_n(dx)$$

by weak convergence. Since $\int_{-\infty}^{\infty} f_m(x)F_n(dx) \leq \int_{-N}^{N} x^2 F_n(dx)$ it follows that

$$\int_{-N}^{N} x^2 F(dx) \leq \lim_{m\to\infty} \limsup_{n\to\infty} \int_{-N}^{N} x^2 F_n(dx) = \limsup_{n\to\infty} \int_{-N}^{N} x^2 F_n(dx)$$

as desired.

18.17 Following the hint, suppose there exists a continuity point y of F such that

$$\lim_{n\to\infty} F_n(y) \neq F(y)$$

Then there exist $\epsilon > 0$ and a subsequence $(n_k)_{k\geq 1}$ s.t. $F_{n_k}(y) - F(y) < -\epsilon$ for all k, or $F_{n_k}(y) - F(y) > \epsilon$ for all k. Suppose $F_{n_k}(y) - F(y) < -\epsilon$ for all k, observe that for $x \leq y$, $F_{n_k}(x) - F(x) \leq F_{n_k}(y) - F(x) = F_{n_k}(y) - F(y) + (F(y) - F(x)) < -\epsilon + (F(y) - F(x))$. Since f is continuous at y there exists an interval $[y_1, y)$ s.t. $|(F(y) - F(x))| < \frac{\epsilon}{2}$, hence $F_{n_k}(x) - F(x) < -\frac{\epsilon}{2}$ for all $x \in [y_1, y)$. Now suppose $F_{n_k}(y) - F(y) > \epsilon$, then for $x \geq y$, $F_{n_k}(x) - F(x) \geq F_{n_k}(y) - F(x) = F_{n_k}(y) - F(y) + (F(y) - F(x)) > \epsilon + (F(y) - F(x))$.

Now we can find an interval $(y, y_1]$ s.t. $|(F(y) - F(x))| < \frac{\epsilon}{2}$ which gives $F_{n_k}(x) - F(x) > \frac{\epsilon}{2}$ for all $x \in (y, y_1]$. Note that both cases would yield

$$\int_{-infty}^{\infty} |F_{n_k}(x) - F(x)|^r dx > |y_1 - y|\frac{\epsilon}{2}$$

which is a contradiction to the assumption

$$\lim_{n\to\infty} \int_{-infty}^{\infty} |F_n(x) - F(x)|^r dx = 0.$$

Therefore X_n converges to X in distribution.

Solutions to selected problems of Chapter 19

19.1 Note that $\varphi_{X_n}(u) = e^{iu\mu_n - \frac{u^2\sigma_n^2}{2}} \to e^{iu\mu - \frac{u^2\sigma^2}{2}}$. By Lévy's continuity theorem it follows that $X_n \Rightarrow X$ where X is $N(\mu, \sigma^2)$.

19.3 Note that $\varphi_{X_n+Y_n}(u) = \varphi_{X_n}(u)\varphi_{Y_n}(u) \to \varphi_X(u)\varphi_Y(u) = \varphi_{X+Y}(u)$. Therefore $X_n + Y_n \Rightarrow X + Y$

Solutions to selected problems of Chapter 20

20.1 a. First observe that $E(S_n^2) = \sum_{i=1}^n \sum_{j=1}^n E(X_i X_j) = \sum_{i=1}^n X_i^2$ since $E(X_i X_j) = 0$ for $i \neq j$. Now $P(\frac{|S_n|}{n} \geq \epsilon) \leq \frac{E(S_n^2)}{\epsilon^2 n^2} = \frac{nE(X_i^2)}{\epsilon^2 n^2} \leq \frac{c}{n\epsilon^2}$ as desired.

b. From part (a) it is clear that $\frac{1}{n}S_n$ converges to 0 in probability. Also $E((\frac{1}{n}S_n)^2) = \frac{E(X_i^2}{n} \to 0$ since $E(X_i^2) \leq \infty$, so $\frac{1}{n}S_n$ converges to 0 in L^2 as well.

20.5 Note that $Z_n \Rightarrow Z$ implies that $\varphi_{Z_n}(u) \to \varphi_Z(u)$ uniformly on compact subset of $\mathbf{R}$. (See Remark 19.1). For any u, we can pick $n > N$ s.t. $\frac{u}{\sqrt{n}} < M$, $\sup_{x\in[-M,M]} |\varphi_{Z_n}(x) - \varphi_Z(x)| < \epsilon$ and $|varphi_Z(\frac{u}{\sqrt{n}}) - \varphi_Z(0)| < \epsilon$. This gives us

$$|\varphi_{Z_n}(\frac{u}{\sqrt{n}}) - \varphi_Z(0)| = |\varphi_{Z_n}(\frac{u}{\sqrt{n}}) - \varphi_Z(\frac{u}{\sqrt{n}})| + |\varphi_Z(\frac{u}{\sqrt{n}}) - \varphi_Z(0)| \leq 2\epsilon$$

So $\varphi_{\frac{Z_n}{\sqrt{n}}}(u) = \varphi_{Z_n}(\frac{u}{\sqrt{n}})$ converges to $\varphi_Z(0) = 1$ for every u. Therefore $\frac{Z_n}{\sqrt{n}} \Rightarrow 0$ by continuity theorem. We also have by the strong law of large numbers that $\frac{Z_n}{\sqrt{n}} \to E(X_j) - \nu$. This implies $E(X_j) - \nu = 0$, hence the assertion follows by strong law of large numbers.

Springer
http://www.springer.com/978-3-540-43871-7

Probability Essentials
Jacod, J.; Protter, P.
2004, X, 254 p., Softcover
ISBN: 978-3-540-43871-7