All of Statistics - Chapter 3 Solutions

May 3, 2020

1.

Let X_n be the number of dollars at the n-th trial. Then,

$$\mathbb{E}[X_{n+1} \mid X_n] = \frac{1}{2}\left(2X_n + \frac{1}{2}X_n\right) = \frac{5}{4}X_n.$$

By the rule of iterated expectations, $\mathbb{E}X_{n+1} = (5/4)\mathbb{E}X_n$. By induction, $\mathbb{E}X_n = (5/4)^n c$.

2.

If $\mathbb{P}(X = c) = 1$, then $\mathbb{E}[X^2] = (\mathbb{E}X)^2 = c^2$ and hence $\mathbb{V}(X) = 0$.

The converse is more complicated. We claim that whenever Y is a nonnegative random variable, $\mathbb{E}Y = 0$ implies that $\mathbb{P}(Y = 0) = 1$. In this case, it is sufficient to take $Y = (X - \mathbb{E}X)^2$ to conclude that $\mathbb{P}(X = \mathbb{E}X) = 1$.

To substantiate the claim, suppose $\mathbb{E}Y = 0$. Take $A_n = \{Y \geq 1/n\}$. Then,

$$0 = \mathbb{E}Y = \mathbb{E}[YI_{A_n} + YI_{A_n^c}] \geq \mathbb{E}[YI_{A_n}] \geq \frac{1}{n}\mathbb{P}(A_n).$$

It follows that $\mathbb{P}(A_n) = 0$ for all n. By continuity of probability,

$$\mathbb{P}(Y > 0) = \mathbb{P}(\cup_n A_n) = \lim_n \mathbb{P}(A_n) = 0.$$

3.

Since $F_{Y_n}(y) = \mathbb{P}(X_1 \leq y)^n = y^n$, it follows that $f_{Y_n}(y) = ny^{n-1}$. Therefore,

$$\mathbb{E}Y_n = n\int_0^1 y^n dy = \frac{n}{n+1}.$$

4.

Note that $X_n = \sum_{i=1}^n (1 - 2B_i) = n - 2\sum_i B_i$ where $B_1, \ldots, B_n \sim \text{Bernoulli}(p)$ are IID. It follows that $\mathbb{E}X_n = n - 2n\mathbb{E}B_1 = n - 2np$ and $\mathbb{V}(X_n) = 4n\mathbb{V}(B_1) = 4np(1-p)$.

5.

Let τ be the number of tosses until a heads is observed. Let C denote the result of the first toss. Then,

$$\mathbb{E}\tau = \frac{1}{2}(\mathbb{E}[\tau \mid C = H] + \mathbb{E}[\tau \mid C = T]) = \frac{1}{2}(1 + (1 + \mathbb{E}\tau))$$

Solving for $\mathbb{E}\tau$ yields 2.

6.

$$\begin{aligned}\mathbb{E}[Y] &= \sum_y y\mathbb{P}(Y = y) = \sum_y y\mathbb{P}(r(X) = y) = \sum_y y\mathbb{P}(X \in r^{-1}(y)) \\ &= \sum_y y \sum_{x \in r^{-1}(y)} \mathbb{P}(X = x) = \sum_y \sum_{x \in r^{-1}(y)} r(x)\mathbb{P}(X = x) = \sum_x r(x)\mathbb{P}(X = x)\end{aligned}$$

7.

Integration by parts yields

$$\begin{aligned}\mathbb{E}X &= \int_0^\infty x f_X(x) dx = \lim_{y \to \infty} yF_X(y) - \int_0^y F_X(x) dx \\ &= \lim_{y \to \infty} \int_0^y F_X(y) - F_X(x) dx = \lim_{y \to \infty} \int_0^\infty (F_X(y) - F_X(x)) I_{(0,y)}(x) dx.\end{aligned}$$

Define $G_y(x) = (F_X(y) - F_X(x))I_{(0,y)}(x)$. Note that G_y converges pointwise to $1 - F_X$ as $y \to \infty$. Moreover, $y \mapsto G_y$ is monotone increasing. The desired result follows by Lebesgue's monotone convergence theorem.

8.

The first two claims follow from

$$\mathbb{E}\overline{X} = \frac{1}{n} \sum_i \mathbb{E}X_i = \mathbb{E}X_1 \equiv \mu$$

and

$$\mathbb{V}(\overline{X}) = \frac{1}{n^2} \sum_i \mathbb{V}(X_1) = \frac{1}{n}\mathbb{V}(X_1) \equiv \frac{\sigma^2}{n}.$$

As for the final claim, note that

$$(n-1) S_n^2 = \sum_i \left(X_i - \overline{X}\right)^2 = \sum_i X_i^2 - 2X_i\overline{X} + \overline{X}^2$$

and hence

$$\frac{n-1}{n}\mathbb{E}\left[S_n^2\right] = \mathbb{E}\left[X_1^2\right] - 2\mathbb{E}\left[X_1\overline{X}\right] + \mathbb{E}\left[\overline{X}^2\right].$$

Next, note that $\mathbb{E}[X_1^2] = \sigma^2 + \mu^2$ and $\mathbb{E}[\overline{X}^2] = \sigma^2/n + \mu^2$. Moreover,

$$X_1\overline{X} = \frac{1}{n}\left(X_1^2 + X_1\sum_{j\neq 1} X_j\right)$$

and hence $\mathbb{E}[X_1\overline{X}] = \sigma^2/n + \mu^2$. Substituting these findings into the equation above yields $\mathbb{E}[S_n^2] = \sigma^2$, as desired.

9.

TODO (Computer Experiment)

10.

The MGF of a normal random variable is $\exp(t^2/2)$. Therefore, $\mathbb{E}\exp(X) = \sqrt{e}$ and

$$\mathbb{V}(\exp(X)) = \mathbb{E}[\exp(2X)] - (\mathbb{E}\exp(X))^2 = e^2 - e.$$

11.

a)

This was already solved in Question 4.

b)

TODO (Computer Experiment)

12.

TODO

13.

a)

Let C denote the result of the coin toss. Then,

$$\mathbb{E}X = \mathbb{E}\left[\mathrm{Unif}(0,1)I_{\{C=H\}} + \mathrm{Unif}(3,4)I_{\{C=T\}}\right] = \frac{1}{2}(\mathbb{E}\,\mathrm{Unif}(0,1) + \mathbb{E}\,\mathrm{Unif}(3,4)) = 2.$$

b)

Similarly to Part (a),

$$\mathbb{E}\left[X^2\right] = \frac{1}{2}\left(\mathbb{E}\left[\text{Unif}(0,1)^2\right] + \mathbb{E}\left[\text{Unif}(3,4)^2\right]\right) = \frac{19}{3}.$$

Therefore, $\mathbb{V}(X) = 19/3 - 4 = 7/3$.

14.

The result follows from

$$\begin{aligned}
&\text{Cov}\left(\sum_i a_i X_i, \sum_j b_j Y_j\right) \\
&\quad = \mathbb{E}\left[\left(\sum_i a_i X_i\right)\left(\sum_j b_j Y_j\right)\right] - \mathbb{E}\left[\sum_i a_i X_i\right]\mathbb{E}\left[\sum_j b_j Y_j\right] \\
&\quad = \sum_{i,j} a_i b_j \mathbb{E}\left[X_i Y_j\right] - \sum_{i,j} a_i b_j \mathbb{E}X_i \mathbb{E}Y_j = \sum_{i,j} a_i b_j \left(\mathbb{E}\left[X_i Y_j\right] - \mathbb{E}X_i \mathbb{E}Y_j\right).
\end{aligned}$$

15.

First, note that $\mathbb{V}(2X - 3Y + 8) = \mathbb{V}(2X - 3Y)$. Moreover,

$$\mathbb{E}\left[(2X - 3Y)^2\right] = \int_0^2 \int_0^1 (2x - 3y)^2 \frac{1}{3}(x + y)\, dx dy = \frac{86}{9}$$

and

$$\mathbb{E}\left[2X - 3Y\right] = \int_0^2 \int_0^1 (2x - 3y)\, \frac{1}{3}(x + y)\, dx dy = -\frac{23}{9}.$$

Therefore, $\mathbb{V}(2X - 3Y) = 245/81$.

16.

In the (absolutely) continuous case,

$$\begin{aligned}
\mathbb{E}\left[r(X)s(Y) \mid X = x\right] = \int r(x)s(y)f_{Y|X}(y \mid x)dy &= r(x)\int s(y)f_{Y|X}(y \mid x)dy \\
&= r(x)\mathbb{E}\left[s(Y) \mid X = x\right].
\end{aligned}$$

Taking $s = 1$ yields $\mathbb{E}[r(X) \mid X = x] = r(x)$. The discrete case is similar. A more general notion of conditional expectation requires Radon-Nikodym derivatives.

17.

By the tower property,

$$\mathbb{E}\left[\mathbb{V}(Y \mid X)\right] = \mathbb{E}\left[\mathbb{E}\left[Y^2 \mid X\right] - \mathbb{E}[Y \mid X]^2\right] = \mathbb{E}\left[Y^2\right] - \mathbb{E}\left[\mathbb{E}[Y \mid X]^2\right]$$

and

$$\mathbb{V}(\mathbb{E}\left[Y \mid X\right]) = \mathbb{E}\left[\mathbb{E}[Y \mid X]^2\right] - \mathbb{E}[\mathbb{E}\left[Y \mid X\right]]^2 = \mathbb{E}\left[\mathbb{E}[Y \mid X]^2\right] - \mathbb{E}[Y]^2.$$

The desired result follows from summing the two quantities.

18.

Since

$$\mathbb{E}[XY] = \mathbb{E}[\mathbb{E}[XY \mid Y]] = \mathbb{E}[\mathbb{E}[X \mid Y]Y] = \mathbb{E}[cY] = c\mathbb{E}Y$$

and $\mathbb{E}X = \mathbb{E}[\mathbb{E}[X \mid Y]] = c$ by the tower property, $\operatorname{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}X\mathbb{E}Y = 0$.

19.

Unlike the distribution of $X_1 \sim \operatorname{Unif}(0, 1)$, the distribution of $(X_1 + \cdots + X_n)/n$ is concentrated around $\mathbb{E}[X_1]$. As n increases, so too does the concentration.

20.

For a vector a with entries a_i,

$$\mathbb{E}\left[a^\intercal X\right] = \mathbb{E}\left[\sum_j a_j X_j\right] = \sum a_j \mathbb{E}X_j = a^\intercal \mathbb{E}X.$$

For a matrix A with entries a_{ij}, define the column vector $a_{i\star}$ as the transpose of the i-th row of A. Then,

$$(\mathbb{E}\left[AX\right])_i = \mathbb{E}\left[(AX)_i\right] = \mathbb{E}\left[a_{i\star}^\intercal X\right] = a_{i\star}^\intercal \mathbb{E}X.$$

Therefore, $\mathbb{E}[AX] = A\mathbb{E}X$.

Next, using our findings in Question 14,

$$\mathbb{V}(a^\intercal X) = \mathbb{V}\left(\sum_j a_j X_j\right) = \sum_{i,j} a_i \operatorname{Cov}(X_i, X_j) a_j = a^\intercal \mathbb{V}(X) a.$$

As before, we can generalize this to the matrix case by noting that

$$(\mathbb{V}(AX))_{ij} = \operatorname{Cov}((AX)_i, (AX)_j) = \operatorname{Cov}(a_{i\star}^\intercal X, a_{j\star}^\intercal X) = \sum_{k,\ell} a_{ik} \operatorname{Cov}(X_k, X_\ell) a_{j\ell}.$$

Therefore, $\mathbb{V}(AX) = A\mathbb{V}(X)A^\intercal$.

21.

If $\mathbb{E}[Y \mid X] = X$, then

$$\mathbb{E}[XY] = \mathbb{E}[\mathbb{E}[XY \mid X]] = \mathbb{E}[X\mathbb{E}[Y \mid X]] = \mathbb{E}[X^2]$$

and $\mathbb{E}Y = \mathbb{E}[\mathbb{E}[Y \mid X]] = \mathbb{E}X$. Therefore,

$$\operatorname{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}X\mathbb{E}Y = \mathbb{E}[X^2] - (\mathbb{E}X)^2 = \mathbb{V}(X).$$

22.

a)

Note that $\mathbb{E}[YZ] = \mathbb{E}I_{(a,b)}(X) = b - a$. Moreover, $\mathbb{E}Y = \mathbb{E}I_{(0,b)}(X) = b$ and $\mathbb{E}Z = \mathbb{E}I_{(a,1)}(X) = 1 - a$. Since $\mathbb{E}[YZ] \neq \mathbb{E}Y\mathbb{E}Z$, Y and Z are dependent.

b)

If $Z = 0$, then $X \leq a < b$ and hence $Y = 1$. Therefore, $\mathbb{E}[Y \mid Z = 0] = 1$ trivially. Moreover,

$$\mathbb{E}\left[Y \mid Z = 1\right] = \frac{\mathbb{E}\left[YZ\right]}{\mathbb{P}(Z = 1)} = \frac{b - a}{1 - a}.$$

23.

Let $K \sim \operatorname{Poisson}(\lambda)$. The MGF of K is

$$\mathbb{E}\left[e^{tK}\right] = e^{-\lambda}\sum_k \frac{\lambda^k e^{tk}}{k!} = e^{-\lambda}\sum_k \frac{\left(\lambda e^t\right)^k}{k!} = \exp(\lambda\left(e^t - 1\right))$$

Let $X \sim N(\mu, \sigma^2)$. Then,

$$\begin{aligned}\sigma\sqrt{2\pi}\mathbb{E}\left[e^{tX}\right] &= \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2\sigma^2}\left((x-\mu)^2 - 2t\sigma^2 x\right)\right\}dx \\ &= \exp\left(t\mu + t^2\sigma^2/2\right)\int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2\sigma^2}\left(x - \mu - t\sigma^2\right)^2\right\}dx.\end{aligned}$$

Therefore, the MGF of X is $\exp(t\mu + t^2\sigma^2/2)$.

Lastly, let $Y \sim \operatorname{Gamma}(\alpha, \beta)$. Then,

$$\mathbb{E}\left[e^{tY}\right] = \beta^{\alpha}\int_0^{\infty}\frac{x^{\alpha-1}e^{(t-\beta)x}}{\Gamma(a)}dx = \left(\frac{\beta}{t-\beta}\right)^{\alpha}\int_0^{\infty}\frac{(t-\beta)^{\alpha}x^{\alpha-1}e^{(t-\beta)x}}{\Gamma(a)}dx$$

is finite whenever $t < \beta$. Therefore, under the same condition, the MGF of Y is $(1-t/\beta)^{-\alpha}$.

24.

Suppose $\beta > t$. Then,

$$\mathbb{E}\left[\exp(tX_1)\right] = \int_0^{\infty}\beta\exp((t-\beta)\,x)dx = \frac{\beta}{\beta-t}$$

and hence

$$\mathbb{E}\left[\exp\left(t\sum_i X_i\right)\right] = \mathbb{E}[\exp(tX_1)]^n = \left(\frac{\beta}{\beta-t}\right)^n = \left(1-\frac{t}{\beta}\right)^{-n}.$$

Since this is the MGF of a Gamma distribution, it follows that the sum of IID exponentially distributed random variables are Gamma distributed.