Exercise Guide for *Statistical Inference (2nd Ed.)* by Casella and Berger

Muthu Chidambaram

Last Updated: July 11, 2019

Contents

About

> *"If you torture the data enough, nature will always confess."* - Ronald Coase

There are actual solutions available here. What follows are notes I took as I worked through some of the exercises in the book on my own.

1 Probability Theory

1.1 Exercise 1

(a) Four character strings consisting of H and T (16 total).

(b) $\mathbb{N}$

(c) $\mathbb{R}^+$

(d) $\mathbb{R}^+/0$

(e) $\frac{i}{n}$ for $i = 0, ..., n$.

1.2 Exercise 4

(a) $P(A) + P(B) - P(A \cap B)$

(b) $P(A) + P(B) - 2P(A \cap B)$

(c) Same as (a).

(d) $1 - P(A \cap B)$

1.3 Exercise 6

We have $u + w = 1 \implies u^2 + 2uw + w^2 = 1$. However, it is not possible for $u^2, 2uw, w^2$ to all equal $\frac{1}{3}$; thus, there are no such u, w.

1.4 Exercise 12

(a) Just a case of countable additivity.

(b) Split $\cup_{i=1}^{\infty} A_i$ into $\cup_{i=1}^{n} A_i$ and $\cup_{i=n+1}^{\infty} A_i$. Taking $n \to \infty$ sends the probability of the latter union to 0, leaving the desired result.

1.5 Exercise 13

No, since $P(B^c) = 1 - P(B) \implies P(B) = \frac{3}{4}$ and $\frac{3}{4} + \frac{1}{3} > 1$.

1.6 Exercise 18

This exercise (based on the provided answer) seems to assume that the balls are distinguishable. There are n^n different arrangements of balls in cells (each ball has n options). If exactly one cell must remain empty, then exactly one other cell must have two balls. There are n choices for the empty cell and $n-1$ choices for the double cell. The two balls to go into the double cell can be chosen in $\binom{n}{2}$ ways, and the remaining balls can be ordered in $(n-2)!$ ways (since they are distinguishable), hence giving the provided result.

1.7 Exercise 21

There are $\binom{2n}{2r}$ ways to choose $2r$ shoes from the collection. If we want to ensure that no pairs of shoes are drawn, we should only draw one shoe from each of the n pairs, which we can do in $\binom{n}{2r}$ ways. For each of the $2r$ shoes we can either choose the right or left shoe from the pair (assuming distinguishability).

1.8 Exercise 23

We use Vandermonde's identity with $m = n = r$ to get

$$\sum_{k=0}^{n} \left(\frac{1}{2^n} \binom{n}{k} \right)^2 = \frac{1}{4^n} \binom{2n}{n}$$

1.9 Exercise 25

Depends on how you define the sample space, see Wikipedia for a lengthy discussion.

1.10 Exercise 26

We take the complement of the probability that we roll a 6 in the first 5 rolls: $1 - \sum_{k=0}^{4} \frac{5^k}{6^{k+1}}$.

1.11 Exercise 35

From the definition of conditional probability and the fact that $P(B) > 0$, we have that $P(\cdot|B) > 0$ and $P(S|B) = \frac{P(B)}{P(B)} = 1$. Countable additivity follows from $A \cap B$ and $A' \cap B$ being disjoint if A and A' are.

1.12 Exercise 36

To solve the first part, we take the complement of the probability that the target was hit fewer than two times, which is

$$1 - \left(\frac{4}{5}\right)^{10} - 10\left(\frac{4}{5}\right)^{9} \frac{1}{5} \approx 0.624$$

If we let A be the event that the target was hit at least once, and B be the event that the target was hit at least twice, then we see that $A \cap B = B$. The conditional probability is thus

$$\frac{1 - \left(\frac{4}{5}\right)^{10} - 10\left(\frac{4}{5}\right)^{9} \frac{1}{5}}{1 - \left(\frac{4}{5}\right)^{10}} \approx 0.699$$

2 Transformations and Expectations

2.1 Exercise 1

(a) $F_Y(y) = P(X^3 \leq y) = F_X(y^{\frac{1}{3}})$. Differentiating, we get that $f_Y(y) = 14y(1 - y^{\frac{1}{3}})$ which integrates to 1 over $(0, 1)$.

(b) $f_Y(y) = \frac{\mathrm{d}}{\mathrm{d}y} F_X(\frac{y-3}{4}) = \frac{7}{4}\exp\left(\frac{-7(y-3)}{4}\right)$, which again integrates to 1 over $(3, \infty)$. The sample space of Y consists of $(3, \infty)$ because $4X+3$ is monotonically increasing and $4(0) + 3 = 3$.

(c) We need only consider positive square roots, since $x \in (0, 1)$. Thus, $f_Y(y) = \frac{1}{2\sqrt{y}} 30y(1 - \sqrt{y})^2$, which integrates to 1 over $(0, 1)$.

2.2 Exercise 2

(a) $f_Y(y) = \frac{1}{2\sqrt{y}} f_X(\sqrt{y}) = \frac{1}{2\sqrt{y}}$

(b) Plug $x = e^{-y}$ into $f_X(x)$ and multiply by e^{-y} (the negative of the derivative, since $-\log X$ is decreasing).

(c) Plug in $x = \log y$ and multiply by $\frac{1}{y}$.

2.3 Exercise 3

From the definition of Y, we can see that the sample space is $\mathcal{Y} = \left(\frac{n}{n+1}\right)_{\mathbb{N}}$. Solving $y = \frac{x}{x+1}$ for y, we find that $f_Y(y) = f_X(\frac{y}{1-y})$.

2.4 Exercise 4

(a) Integrating from $-\infty$ to 0 gives $\frac{1}{2}$ and likewise for 0 to ∞, so f is a pdf.

(b) If $t \leq 0$, then we have that $P(X < t) = \frac{1}{2}e^{\lambda t}$. Otherwise,

$$\begin{aligned} P(X < t) &= \int_{-\infty}^{t} f(x) = \int_{-\infty}^{0} f(x) + \int_{0}^{t} f(x) \\ &= \frac{1}{2} + \frac{1}{2} - \frac{1}{2}e^{-\lambda t} \\ &= 1 - \frac{1}{2}e^{-\lambda t} \end{aligned}$$

(c) If $t \leq 0$, $P(|X| < t)$ is clearly 0. Otherwise,

$$\begin{aligned} P(|X| < t) = P(-t < X < t) &= P(X < t) - P(X < -t) \\ &= 1 - e^{-\lambda t} \end{aligned}$$

from the CDFs computed in part (b).

2.5 Exercise 5

The sample space $\mathcal{Y}$ is $[0,1]$. Thus, we need to consider $P(Y \leq y) = P(X \leq \sin^{-1}\sqrt{y})$ for $y \in [0,1]$. By symmetry, we can just consider the case where $\sin^{-1}$ is restricted to $[0, \frac{\pi}{2}]$, as the other three quadrants have the same area. Therefore,

$$\begin{aligned} f_Y(y) &= 4\frac{\mathrm{d}}{\mathrm{d}y}F_X(\sin^{-1}\sqrt{y}) \\ &= 4\frac{\mathrm{d}}{\mathrm{d}y}\frac{\sin^{-1}\sqrt{y}}{2\pi} \\ &= \frac{1}{\pi\sqrt{y(1-y)}} \end{aligned}$$

At $y = 0$ and $y = 1$, the density is infinite/undefined - I'm not really sure how to interpret this.

2.6 Exercise 11

(a) We have that

$$\begin{aligned} \mathrm{E}[X^2] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} x^2 e^{-\frac{x^2}{2}} &= \frac{2}{\sqrt{2\pi}}\int_{0}^{\infty} x^2 e^{-\frac{x^2}{2}} \\ &= \frac{2}{\sqrt{2\pi}}\big(\lim_{x\to\infty} -xe^{-\frac{x^2}{2}} + \int_0^{\infty} e^{-\frac{x^2}{2}}\big) \\ &= 1 \end{aligned}$$

Where we integrated by parts using $dv = xe^{-\frac{x^2}{2}}$ and $u = x$, applied L'Hôpital's rule, and then compared to the CDF of the normal distribution.

(b) The pdf $f_Y(y)$ is just $f_X(y) + f_X(-y)$, so $f_Y(y) = \frac{2}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}$. We can then compute $\mathrm{E}[Y]$ as

$$\mathrm{E}[Y] = \frac{2}{\sqrt{2\pi}}\int_0^{\infty} ye^{-\frac{y^2}{2}} = \frac{2}{\sqrt{2\pi}}$$

To compute the variance we need to compute $\mathrm{E}[Y^2]$, which is identical to $\mathrm{E}[X^2]$. Thus, $\mathrm{Var}[Y] = \mathrm{E}[Y^2] - \mathrm{E}[Y]^2 = 1 - \frac{2}{\pi}$.