Probability and Statistics

Exercise sheet 9

Exercise 9.1 (A gambling example).

It costs 1 dollar to play a certain slot machine in Las Vegas. The machine is set by the house to pay 2 dollars with probability 0.45 and to pay nothing with probability 0.55.

Let X_i be the house's net winnings on the i^{th} play of the machine.

Let $S_n := \sum_{i=1}^n X_i$ be the house's winnings after n plays of the machine. Assuming that successive plays are independent, find:

(a) $E(S_n)$;

(b) $\operatorname{var}(S_n)$;

(c) the approximate probability that after 10,000 rounds of the machine, the house's winnings are between 800 and 1,100 dollars ($800 \leq S_{10,000} \leq 1,100$).

Solution 9.1 The house's winnings resulting from each play are independent, and take values

$$X_i = \begin{cases} 1, & \text{with probability } 0.55, \\ -1, & \text{with probability } 0.45. \end{cases}$$

Therefore, one can observe that

$$B_i := \frac{X_i + 1}{2}$$

are independent Ber(0.55) random variables, and that

$$\tilde{S}_n = \frac{S_n + n}{2} \Leftrightarrow S_n = 2\tilde{S}_n - n$$

where $\tilde{S}_n := \sum_{i=1}^n B_i$ has a $\text{Bin}(n, 0.55)$ distribution.

From the considerations above,

(a)
$$E(S_n) = 2E(\tilde{S}_n) - n = 1.1n - n = 0.1n$$

and

(b)
$$\operatorname{var}(S_n) = 4\operatorname{var}(\tilde{S}_n) = 4 \times 0.55 \times 0.45n = 0.99n,$$

by the linearity properties of expectations and variance, and using the known values for a Binomial distribution.

(c) Since the X_n are i.i.d, by the central limit theorem we can find a good Gaussian approximation to this probability. Since we calculated the expectation and variance before, for large n,

$$\frac{S_n - 0.1n}{\sqrt{0.99n}} \overset{\text{d}}{\approx} \mathcal{N}(0,1),$$

and in particular:

$$\frac{S_{10,000} - 1,000}{\sqrt{9,900}} \overset{\text{d}}{\approx} \mathcal{N}(0,1).$$

Therefore we obtain:

$$\begin{aligned}
P(800 \leq S_{10,000} \leq 1,100) &= P\left(-\frac{200}{\sqrt{9,900}} \leq \frac{S_n - 1,000}{\sqrt{9,900}} \leq \frac{100}{\sqrt{9,900}}\right) \\
&\approx \Phi\left(\frac{100}{\sqrt{9,900}}\right) - \Phi\left(-\frac{200}{\sqrt{9,900}}\right) \\
&\approx 0.82.
\end{aligned}$$

Exercise 9.2 Suppose that the cost of a textbook at the college level is on average 50 francs, with a standard deviation of 7 francs.

In a four year bachelor's program, a student will need to buy 25 textbooks, the prices of which are assumed to be independent and identically distributed. Find an approximation to the probability that the student will have to spend more than 1300 francs on textbooks.

Solution 9.2 If the prices of the textbooks are P_i for $i = 1, ..., 25$, the total expenditure is

$$S_{25} = \sum_{i=1}^{25} P_i.$$

We are given that the P_i are independent, with common mean $E(P_i) = 50$ and variance $\text{var}(P_i) = 7^2 = 49$ (i.e. standard deviation $\sigma = 7$).

Assuming that $n = 25$ is large enough for the central limit theorem to give a good approximation, we then obtain

$$\frac{S_{25} - n\mu}{\sqrt{n}\sigma} = \frac{S_{25} - 1,250}{35} \overset{\text{d}}{\approx} \mathcal{N}(0,1)$$

and so we can approximate:

$$\begin{aligned}
P(S_{25} > 1,300) &= P\left(\frac{S_{25} - 1,250}{35} > \frac{50}{35}\right) \\
&\approx 1 - \Phi\left(\frac{10}{7}\right) \\
&\approx 0.0766.
\end{aligned}$$

Exercise 9.3 (Comparing a Poisson approximation and a normal approximation).

Suppose 1.5% of residents of a town never read a newspaper. We draw a random sample of 50 people, and we want to determine the probability that at least 1 resident in the sample never reads a newspaper.

For that probability, compute:

(a) the exact value;

(b) a Poisson approximation;

(c) a normal approximation.

Solution 9.3 Since we do not know the total number of residents of the town (if we did, we could use a Hypergeometric distribution), the next best thing is to assume that the number is large enough that the selected residents can be assumed to be independent from each other (in terms of whether they read a newspaper or not).

Therefore, we model the random sample by $X_1, ..., X_{50} \overset{\text{iid}}{\sim} \text{Ber}(50, 0.015)$, where $X_i = 1$ with probability $p = 1.5\%$ if resident i in the sample never reads a newspaper. Then

$$S_{50} = \sum_{i=1}^{50} X_i \sim \text{Bin}(50, 0.015)$$

is the total number of residents in the sample that never read a newspaper. We want to compute $P(S_n \geq 1) = 1 - P(S_n = 0)$.

(a)
$$1 - P(S_{50} = 0) = 1 - \binom{50}{0} p^0 (1-p)^{50} = 1 - 0.985^{50} \approx 0.5303.$$

(b) Note that
$$S_n \sim \text{Bin}(50, 0.015) = \text{Bin}\left(50, \frac{0.75}{50}\right) \approx \text{Poi}(0.75),$$
by taking the Poisson approximation assuming that 50 is large enough. Under the Poisson approximation,
$$1 - P(S_{50} = 0) \approx 1 - \frac{e^{-0.75} 0.75^0}{0!} = 1 - e^{-0.75} \approx 0.5276.$$

(c) The expectation and variance of a binomial distribution are
$$E(S_{50}) = 50 \times 0.015 = 0.75$$
and
$$\text{var}(S_{50}) = 50 \times 0.015 \times 0.985 = 0.73875.$$
The normal approximation gives
$$\frac{S_{50} - 0.75}{\sqrt{0.73875}} \overset{\text{d}}{\approx} \mathcal{N}(0, 1)$$
and so we can estimate
$$\begin{aligned}
1 - P(S_{50} = 0) &= 1 - P(S_{50} \leq 0) \\
&= 1 - P\left(\frac{S_{50} - 0.75}{\sqrt{0.73875}} \leq \frac{-0.75}{\sqrt{0.73875}}\right) \\
&\approx 1 - \Phi\left(\frac{-0.75}{\sqrt{0.73875}}\right) \\
&\approx 0.81
\end{aligned}$$
As we can see, the normal approximation in this case is quite far off. One explanation is that $n = 50$ is not enough for this distribution to look like a normal distribution, since a lot of the probability ($\sim 47\%$) is still concentrated in a single point $S_{50} = 0$, which should not be the case for an approximately normal random variable.

Exercise 9.4 Consider the joint pmf

$$p(x, y) = \begin{cases} cxy, & 1 \leq x \leq 3, 1 \leq y \leq 3, \\ 0, & \text{otherwise.} \end{cases}$$

(a) Find the normalising constant c.

(b) Are X and Y independent? Why?

(c) Find $E(X)$, $E(Y)$ and $E(XY)$.

Solution 9.4
Remark: The exercise was intended to mean that X, Y are discrete random variables, taking values on $\{1,2,3\} \times \{1,2,3\}$ (version 1). However, one could also consider describing a continuous joint distribution on $[1,3] \times [1,3]$ (version 2). Solutions are given for both versions.

Version 1:

(a) We find the normalising constant by summing:

$$\begin{aligned} 1 &= \sum_{x,y} p(x,y) \\ &= \sum_{x=1}^{3} \sum_{y=1}^{3} cxy \\ &= c \sum_{x=1}^{3} x \sum_{y=1}^{3} y \\ &= c \times 6 \times 6 \\ &= 36c \end{aligned}$$

so $c = \frac{1}{36}$.

(b) Yes, X and Y are independent. We can see this immediately, since the joint pmf can be factorised:

$$p_{X,Y}(x,y) = \frac{xy}{36} \mathbb{1}_{x,y \in \{1,2,3\}} = \left(\frac{x}{6} \mathbb{1}_{\{1,2,3\}} \right) \left(\frac{y}{6} \mathbb{1}_{\{1,2,3\}} \right).$$

To give some more detail, note that the marginal pmf of X is

$$\begin{aligned} p_X(x) &= \sum_{y} p_{X,Y}(x,y) \\ &= \sum_{y=1}^{3} \frac{xy}{36} \mathbb{1}_{x \in \{1,2,3\}} \\ &= \frac{x}{6} \mathbb{1}_{x \in \{1,2,3\}}, \end{aligned}$$

and by a similar calculation

$$p_Y(y) = \frac{y}{6} \mathbb{1}_{y \in \{1,2,3\}},$$

and indeed $p_{X,Y}(x,y) = p_X(x) p_Y(y)$, showing that they are independent.

(c) We compute

$$E(X) = \sum_{x} x p_X(x) = \sum_{x=1}^{3} x \frac{x}{6} = \frac{7}{3}$$

and analogously
$$E(Y) = \frac{7}{3}.$$
By independence,
$$E(XY) = E(X)E(Y) = \left(\frac{7}{3}\right)^2.$$

Version 2:

Here instead of a pmf, we consider the joint density
$$f_{X,Y}(x,y) = \begin{cases} cxy, & 1 \leq x \leq 3, 1 \leq y \leq 3, \\ 0, & \text{otherwise.} \end{cases}$$

(a) We find the normalising constant by integrating:
$$\begin{aligned} 1 &= \int_{\mathbb{R}^2} f_{X,Y}(x,y) dxdy \\ &= \int_1^3 \int_1^3 cxy \, dxdy \\ &= c \int_1^3 xdx \int_1^3 ydy \\ &= c \times 4 \times 4 \\ &= 16c \end{aligned}$$
so $c = \frac{1}{16}$.

(b) Yes, X and Y are independent. We can see this immediately, since the joint density can be factorised:
$$f_{X,Y}(x,y) = \frac{xy}{16} \mathbb{1}_{1 \leq x \leq 3, 1 \leq y \leq 3} = \left(\frac{x}{4} \mathbb{1}_{1 \leq x \leq 3}\right)\left(\frac{y}{4} \mathbb{1}_{1 \leq y \leq 3}\right).$$
To give some more detail, note that the marginal density of X is
$$\begin{aligned} f_X(x) &= \int_{\mathbb{R}} f_{X,Y}(x,y) dy \\ &= \int_1^3 \frac{xy}{16} \mathbb{1}_{1 \leq x \leq 3} dy \\ &= \frac{x}{4} \mathbb{1}_{1 \leq x \leq 3}, \end{aligned}$$
and by a similar calculation
$$f_Y(y) = \frac{y}{4} \mathbb{1}_{1 \leq y \leq 3},$$
and indeed $f_{X,Y}(x,y) = f_X(x) f_Y(y)$, showing that they are independent.

(c) We compute
$$E(X) = \int_{\mathbb{R}} x f_X(x) dx = \int_1^3 x \frac{x}{4} dx = \frac{13}{6}$$

and analogously

$$E(Y) = \frac{13}{6}.$$

By independence,

$$E(XY) = E(X)E(Y) = \left(\frac{13}{6}\right)^2.$$

Exercise 9.5 Let (X, Y) be a discrete random pair.

(a) Show that

$$\text{var}(X) = \text{var}_Y(E(X \mid Y = y)) + E_Y(\text{var}(X \mid Y = y)).$$

(b) Suppose that $Y \sim \text{Poi}(\lambda)$ for some $\lambda \in (0, +\infty)$ and $X \mid Y = y \sim \text{U}(\{0, 1, ..., y\})$. Find $E(X)$ and var(X).

(c) Now, suppose that X and Y are $\overset{\text{iid}}{\sim}$ Geo(p) for some $p \in (0, 1)$. What is $P(X \geq Y)$? And $P(X > Y)$?

Solution 9.5

(a) Since the pair (X, Y) is discrete, it only takes with positive probability a countable collection of values $\{v_1 = (x_1, y_1), v_2 = (x_2, y_2), ...\}$. Thus, X only takes with positive probability the countable collection of values $\{x_1, x_2, ...\}$ and Y only takes with positive probability the countable collection of values $\{y_1, y_2, ...\}$ (some of these could be repeated, and some could have 0 probability, i.e. we could have only finitely many values). In other words, X and Y are individually discrete as well.

Now, we want to show that

$$E(\text{var}(X \mid Y)) + \text{var}(E(X \mid Y)) = \text{var}(X).$$

Here it is assumed that $\text{var}(X) < \infty$, which is equivalent to $E(X^2) < \infty$.

$$\begin{aligned} E(\text{var}(X \mid Y)) &= \sum_y p_Y(y) \sum_x p(x \mid y)(x - E(X \mid Y = y))^2 \\ &= \sum_{x,y} p_Y(y)p(x \mid y)(x - E(X \mid Y = y))^2 \quad \text{(by Fubini, since the terms} \geq 0) \\ &= \sum_{x,y} p_Y(y)p(x \mid y)(x^2 - 2xE(X \mid Y = y) + E(X \mid Y = y)^2). \end{aligned}$$

The first term is

$$\sum_{x,y} p_Y(y)p(x \mid y)x^2 = E(X^2),$$

since by Fubini's theorem,

$$\begin{aligned}\sum_{x,y} p_Y(y)p(x \mid y)x^2 &= \sum_x x^2 \sum_y p_Y(y)p(x \mid y)\\ &= \sum_x x^2 \sum_y p(x,y)\\ &= \sum_x x^2 p_X(x)\\ &= E(X^2).\end{aligned}$$

For the next term, note that also

$$\sum_{x,y} p_Y(y)p(x \mid y)|x||E(X \mid Y=y)| < \infty.$$

To see this, note that

$$\begin{aligned}|E(X \mid Y=y)| &= \left|\sum_x p(x \mid y)x\right|\\ &\leq \sum_x |x|p(x \mid y)\\ &= E(|X| \mid Y=y)\end{aligned}$$

and by Fubini's theorem,

$$\begin{aligned}&\sum_{x,y} p_Y(y)p(x \mid y)|x|E(|X| \mid Y=y)\\ &= \sum_y p_Y(y)E(|X| \mid Y=y)\sum_x |x|p(x \mid y)\\ &= \sum_y p_Y(y)E(|X| \mid Y=y)E(|X| \mid Y=y)\\ &= \sum_y p_Y(y)E(|X| \mid Y=y)^2\\ &= \sum_y p_Y(y)E(X^2 \mid Y=y) \quad \text{(by Jensen's inequality applied to } X \mid Y=y)\\ &= E(X^2) < \infty.\end{aligned}$$

Thus, we are allowed to write, using Fubini:

$$\begin{aligned}\sum_{x,y} p_Y(y)p(x \mid y)xE(X \mid Y=y) &= \sum_y p_Y(y)E(X \mid Y=y)\sum_x p(x \mid y)x\\ &= \sum_y p_Y(y)E(X \mid Y=y)^2.\end{aligned}$$

It follows that

$$E(\text{var}(X \mid Y)) = E(X^2) - 2\sum_y p_Y(y)E(X \mid Y = y)^2 + \sum_{x,y} p_Y(y)p(x \mid y)E(X \mid Y = y)^2$$

where

$$\sum_{x,y} p_Y(y)p(x \mid y)E(X \mid Y = y)^2 = \sum_y p_Y(y)E(X \mid Y = y)^2 \sum_x p(x \mid y)$$
$$= \sum_y p_Y(y)E(X \mid Y = y)^2$$

yielding $E(\text{var}(X \mid Y)) = E(X^2) - \sum_y p_Y(y)E(X \mid Y = y)^2$.
On the other hand,

$$\text{var}(E(X \mid Y)) = E(E(X \mid Y)^2) - E(E(X \mid Y))^2 = E(E(X \mid Y)^2) - E(X)^2$$

by the iterated expectation formula.
But $E(E(X \mid Y)^2) = \sum_y p_Y(y)E(X \mid Y = y)^2$ yielding

$$E(\text{var}(X \mid Y)) + \text{var}(E(X \mid Y)) = E(X^2) - E(X^2) = \text{var}(X),$$

as we wanted.

(b) Conditioning on Y, and since we know the expectation and variance of a uniform random variable,

$$\begin{aligned} E(X) &= \sum_{y=0}^{\infty} \frac{\lambda^y e^{-\lambda}}{y!} E(X \mid Y = y) \\ &= \sum_{y=0}^{\infty} \frac{\lambda^y e^{-\lambda}}{y!} \frac{y}{2} \\ &= \frac{1}{2} \sum_{y=0}^{\infty} y \frac{\lambda^y e^{-\lambda}}{y!} \\ &= \frac{1}{2} E(Y) = \frac{\lambda}{2}, \end{aligned}$$

$$\begin{aligned} \text{var}(X) &= \text{var}_Y(E(X \mid Y = y)) + E_Y(\text{var}(X \mid Y = y)) \\ &= \text{var}\left(\frac{Y}{2}\right) + E\left(\frac{(Y+1)^2 - 1}{12}\right) \\ &= \frac{\lambda}{4} + \frac{\lambda + \lambda^2 + 2\lambda}{12} \\ &= \frac{\lambda}{2} + \frac{\lambda^2}{12}. \end{aligned}$$

(c) We can compute

$$\begin{aligned} P(X=Y) &= \sum_{k=1}^{\infty} P(X=Y=k) \\ &= \sum_{k=1}^{\infty} P(X=k, Y=k) \\ &= \sum_{k=1}^{\infty} P(X=k)P(Y=k) \\ &= \sum_{k=1}^{\infty} p^2(1-p)^{2k-2} \\ &= \frac{p^2}{1-(1-p)^2} \\ &= \frac{p}{2-p}. \end{aligned}$$

Note that by the symmetry of the problem, $P(X<Y) = P(X>Y)$ and so

$$P(X<Y) + P(X=Y) + P(X>Y) = 1 \Rightarrow \; P(X>Y) = \frac{1-\frac{p}{2-p}}{2} = \frac{1-p}{2-p}$$

and

$$P(X \geq Y) = P(X>Y) + P(X=Y) = \frac{1}{2-p}.$$