Chapter 3 Solutions

Andrew Wu
Wasserman: All of Statistics

February 17, 2025

Problem 3.1. Suppose we play a game where you start with c dollars. On each turn of the game you either halve or double your money, with equal probability. What is your expected fortune after n turns?

Solution. Let X be a random variable denoting how much money you have after one turn of the game if you originally have a dollars.

Then $\mathbb{E}(X) = \sum xf(x)$, where $f(x)$ is the PMF of X. We have $f(\frac{a}{2}) = f(2a) = \frac{1}{2}$, and $f(x) = 0$ for all other x.

So

$$\begin{aligned}\mathbb{E}(X) &= 2a \cdot f(2a) + \frac{a}{2} \cdot f(\frac{a}{2}) \\ &= a + \frac{a}{4} \\ &= \frac{5a}{4}.\end{aligned}$$

Note this is true for any a, so your expected fortune after n turns is $c \cdot (\frac{5}{4})^n$. □

Problem 3.2. Show that $\mathbb{V}(X) = 0$ if and only if there is a constant c such that $\mathbb{P}(X = c) = 1$.

Solution. By definition, we have

$$\mathbb{V}(X) = \mathbb{E}((X - \mathbb{E}(X))^2).$$

If there is a constant c such that $\mathbb{P}(X = c) = 1$, then $X = c$ and

$$\begin{aligned}\mathbb{V}(X) &= \mathbb{E}((X - c)^2) \\ &= \mathbb{E}(0) = 0.\end{aligned}$$

On the other hand, if $\mathbb{V}(x) = 0$, then $\mathbb{E}((X - \mathbb{E}(X))^2) = 0$. Note that by the Trivial Inequality, $(X - \mathbb{E}(X))^2$ is nonnegative, so for the equation $\mathbb{E}((X - \mathbb{E}(X))^2 = 0$ to hold, we must have $X - \mathbb{E}(X) = 0$ for all possible values of X. Thus $X = \mathbb{E}(X)$, so X must be equal to some constant, or equivalently, $\mathbb{P}(X = c) = 1$ for some constant c. □

Problem 3.3. Let $X_1, \ldots, X_n \sim \text{Uniform}(0, 1)$ and let $Y_n = \max(X_1, \ldots, X_n)$. Find $\mathbb{E}(Y_n)$.

Solution. We have

$$\mathbb{E}(Y_n) = \int y f_Y(y) dy,$$

where $f_Y(y)$ is the PDF of Y_n.

We begin by finding $F_Y(y)$, the CDF of Y_n. We have

$$\begin{aligned}F_Y(y) &= \mathbb{P}(Y_n \leq y) \\ &= \mathbb{P}(X_1, \ldots, X_n \leq y) \\ &= \mathbb{P}(X_1 \leq y) \cdot \ldots \cdot \mathbb{P}(X_n \leq y) \\ &= y^n.\end{aligned}$$

Thus $f_Y(y) = ny^{n-1}$, and

$$\begin{aligned}\mathbb{E}(Y_n) &= \int y f_Y(y) dy \\ &= \int_0^1 ny^n dy \\ &= \frac{n}{n-1}.\end{aligned}$$

□

Problem 3.4. A particle stands at the origin of the real line and moves along the line in jumps of one unit. For each jump the probability is p that the particle will jump one unit to the left and the probability is $1-p$ that the particle will jump one unit to the right. Let X_n be the position of the particle after n jumps. Find $\mathbb{E}(X_n)$ and $\mathbb{V}(X_n)$. (This is known as a random walk.)

Solution. Note that X_n can be written as the sum $Y_1 + Y_2 + \ldots + Y_n$, where each Y_i is a random variable equal to -1 with probability p and equal to 1 with probability $1-p$, and where all the Y_is are independent.

We have

$$\begin{aligned}\mathbb{E}(Y_i) &= (-1) \cdot p + 1 \cdot (1-p) \\ &= 1 - 2p\end{aligned}$$

so therefore

$$\begin{aligned}\mathbb{E}(X_n) &= \mathbb{E}(Y_1 + \ldots + Y_n) \\ &= \mathbb{E}(Y_1) + \ldots + \mathbb{E}(Y_n) \\ &= n(1-2p).\end{aligned}$$

Next, observe that as the Y_i are independent, we have

$$\begin{aligned}\mathbb{V}(X_n) &= \mathbb{V}(Y_1 + \ldots + Y_n) \\ &= \mathbb{V}(Y_1) + \ldots + \mathbb{V}(Y_n) \\ &= n\mathbb{V}(Y_i).\end{aligned}$$

Thus we need only compute $\mathbb{V}(Y_i)$. We have

$$\begin{aligned}\mathbb{V}(Y_i) &= \mathbb{E}(Y_i^2) - \mathbb{E}(Y_i)^2 \\ &= 1 - (1-2p)^2 \\ &= 4p - 4p^2\end{aligned}$$

where we used that $Y_i^2 = 1$ because $Y_i = 1$ or $Y_i = -1$.

Thus our answer is $\mathbb{V}(X_n) = n(4p - 4p^2)$. □

Problem 3.5. A fair coin is tossed until a head is obtained. What is the expected number of tosses that will be required?

Solution. Let X denote the number of tosses necessary to obtain a head, with PMF $f(x)$.

Observe that $f(n) = (\frac{1}{2})^n$, as one must toss $n-1$ tails in a row, with probability $(\frac{1}{2})^{n-1}$, and then a head, with probability $\frac{1}{2}$; the product of these probabilities is $(\frac{1}{2})^n$.

Then

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{i=1}^{\infty} x f(x) \\
&= 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{8} + \dots \\
&= \left(\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \dots\right) + \left(\frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \dots\right) + \left(\frac{1}{8} + \frac{1}{16} + \frac{1}{32} + \dots\right) + \dots \\
&= 1 + \frac{1}{2} + \frac{1}{4} + \dots \\
&= 2.
\end{aligned}$$

□

Problem 3.6. Let $Y = r(X)$ for discrete random variables X and Y. Prove that $\mathbb{E}(Y) = \mathbb{E}(r(X)) = \sum r(x) f_X(x)$.

Solution. Let $A_y = \{x : r(x) = y\}$.

We have $f_Y(y) = \sum_{A_y} f_X(x)$, for all y. Then

$$\begin{aligned}
\mathbb{E}(Y) = \sum_y y f_Y(y) &= \sum_y \left(y \sum_{A_y} f_X(x) \right) \\
&= \sum_y \left(y \sum_{x_1, \dots, x_k} f_X(x) \right) \\
&= \sum_x r(x) f_X(x),
\end{aligned}$$

where we used the fact that $r(x_1) = \dots = r(x_k) = y$. □

Problem 3.7. Let X be a continuous random variable with CDF F. Suppose that $\mathbb{P}(X > 0) = 1$ and that $\mathbb{E}(X)$ exists. Show that $\mathbb{E}(X) = \int_0^\infty \mathbb{P}(X > x) dx$.

Solution. We have $\mathbb{E}(X) = \int x f(x) dx$ by definition, and as $\mathbb{P}(X > 0) = 1$, we have

$$\mathbb{E}(X) = \int_0^\infty x f(x) dx.$$

Now we integrate by parts. We have

$$\begin{aligned}
\int_0^\infty x f(x) dx &= [x F(x)]_0^\infty - \int_0^\infty F(x) dx \\
&= \lim_{x \to \infty} x F(x) - \int_0^\infty F(x) dx \\
&= \lim_{x \to \infty} x - \int_0^\infty F(x) dx \\
&= \int_0^\infty 1 dx - \int_0^\infty F(x) dx \\
&= \int_0^\infty 1 - F(x) dx \\
&= \int_0^\infty \mathbb{P}(X > x) dx,
\end{aligned}$$

where we have used the fact that as $\mathbb{E}(X)$ exists, $\lim_{x \to \infty} [x F(x) - x] = 0$. □

Problem 3.8. Let $X_1, \ldots, X_n$ be IID and let $\mu = \mathbb{E}(X_i), \sigma^2 = \mathbb{V}(X_i)$. Prove that

$$\mathbb{E}(\overline{X}_n) = \mu, \mathbb{V}(\overline{X}_n) = \frac{\sigma^2}{n}, \quad \text{and} \quad \mathbb{E}(S_n^2) = \sigma^2.$$

Solution. We have

$$\begin{aligned}
\mathbb{E}(\overline{X}_n) &= \mathbb{E}\left(\frac{X_1 + \ldots + X_n}{n}\right) \\
&= \frac{1}{n}\left(\mathbb{E}(X_1) + \ldots + \mathbb{E}(X_n)\right) \\
&= \frac{1}{n}(n \cdot \mu) = \mu.
\end{aligned}$$

Next,

$$\begin{aligned}
\mathbb{V}(\overline{X}_n) &= \mathbb{V}\left(\frac{X_1 + \ldots + X_n}{n}\right) \\
&= \frac{1}{n^2}(\mathbb{V}(X_1) + \ldots + \mathbb{V}(X_n)) \\
&= \frac{1}{n^2}(n\sigma^2) = \frac{\sigma^2}{n}.
\end{aligned}$$

To solve the last part, we expand the summation:

$$\begin{aligned}
\mathbb{E}(S_n^2) &= \mathbb{E}\left(\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2\right) \\
&= \frac{1}{n-1}\mathbb{E}\left(\sum_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2\right) \\
&= \frac{1}{n-1}\mathbb{E}\left(\sum_{i=1}^{n}\left(X_i^2 + \overline{X}_n^2 - 2X_i\overline{X}_n\right)\right) \\
&= \frac{1}{n-1}\mathbb{E}\left(\sum_{i=1}^{n}X_i^2 + \sum_{i=1}^{n}\overline{X}_n^2 - 2\sum_{i=1}^{n}\left(X_i\overline{X}_n\right)\right) \\
&= \frac{1}{n-1}\mathbb{E}\left(\sum_{i=1}^{n}X_i^2 + n \cdot \overline{X}_n^2 - \frac{2}{n}\sum_{i=1}^{n}(X_i(X_1 + \ldots + X_n))\right) \\
&= \frac{1}{n-1}\mathbb{E}\left(\sum_{i=1}^{n}X_i^2 + \frac{1}{n}(X_1 + \ldots + X_n)^2 - \frac{2}{n}(X_1 + \ldots + X_n)^2\right) \\
&= \frac{1}{n-1}\mathbb{E}\left(\sum_{i=1}^{n}X_i^2 - \frac{1}{n}(X_1 + \ldots + X_n)^2\right) \\
&= \frac{1}{n-1}\mathbb{E}\left(\frac{n-1}{n}\left(X_1^2 + \ldots + X_n^2\right) - \frac{1}{n}\sum_{1 \le i < j \le n} 2X_iX_j\right).
\end{aligned}$$

Now we can use linearity of expectation. Begin by noting that

$$\begin{aligned}
\frac{1}{n-1}\mathbb{E}\left(\frac{n-1}{n}\left(X_1^2 + \ldots + X_n^2\right)\right) &= \frac{1}{n}\mathbb{E}(X_1^2 + \ldots + X_n^2) \\
&= \frac{1}{n}(n \cdot \mathbb{E}(X_i^2)) = \mathbb{E}(X_i^2).
\end{aligned}$$

Next, observe that

$$\begin{aligned}
\frac{1}{n-1}\mathbb{E}\left(\frac{1}{n}\sum_{1\le i<j\le n} 2X_iX_j\right) &= \frac{2}{n(n-1)}\mathbb{E}\left(\sum_{1\le i<j\le n} X_iX_j\right) \\
&= \frac{2}{(n(n-1)} \cdot \frac{n(n-1)}{2} \cdot \mathbb{E}(X_iX_j) \\
&= \mu^2,
\end{aligned}$$

using the fact that there are $\binom{n}{2}$ pairs (i,j) with $1 \le i < j \le n$.

Putting it all together, we obtain

$$\mathbb{E}(S_n^2) = \mathbb{E}(X_i^2) - \mu^2 = \sigma^2,$$

as desired. □

Problem 3.10. Let $X \sim N(0,1)$ and let $Y \sim e^X$. Find $\mathbb{E}(Y)$ and $\mathbb{V}(Y)$.

Solution. Note that e^x is a strictly monotone increasing function. Therefore,

$$f_Y(y) = f_X(s(y))\left|\frac{ds(y)}{dy}\right|$$

where s is the inverse of e^x.

Thus

$$\begin{aligned}
f_Y(y) &= \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}s(y)^2\right)\cdot\left|\frac{ds(y)}{dy}\right| \\
&= \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\ln(y)^2\right)\cdot\left|\frac{1}{y}\right| \\
&= \frac{1}{y\sqrt{2\pi}}\exp\left(-\frac{1}{2}\ln(y)^2\right)
\end{aligned}$$

when $y > 0$, using the fact that Y must be positive.

Now we want to compute

$$\begin{aligned}
\mathbb{E}(Y) &= \int y f_Y(y)dy \\
&= \int_0^\infty \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\ln(y)^2\right)dy \\
&= \frac{1}{\sqrt{2\pi}}\int_0^\infty \exp\left(-\frac{1}{2}\ln(y)^2\right)dy.
\end{aligned}$$

A u-substitution with $\ln y = u$ yields

$$\frac{1}{\sqrt{2\pi}}\int_0^\infty \exp\left(-\frac{1}{2}\ln(y)^2\right)dy = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^\infty \exp\left(-\frac{1}{2}u^2 - u\right)du.$$

(Note that one can obtain the right-hand-side directly through the Law of the Unconscious Statistician.) We'll evaluate the integral by itself and multiply back by $\frac{1}{\sqrt{2\pi}}$ later, using the method of completing the square. Observe that

$$\begin{aligned}
\int_{-\infty}^\infty \exp\left(-\frac{1}{2}u^2 - u\right)du &= e^{\frac{1}{2}}\int_{-\infty}^\infty \exp\left(-\frac{1}{2}u^2 - u - \frac{1}{2}\right)du \\
&= e^{\frac{1}{2}}\int_{-\infty}^\infty \exp\left(-\left(\frac{1}{\sqrt{2}}u - \frac{1}{\sqrt{2}}\right)^2\right)du \\
&= \sqrt{2e}\int_{-\infty}^\infty \exp(-v^2)dv
\end{aligned}$$

where the last step follows from substituting $v = \frac{1}{\sqrt{2}}u + \frac{1}{\sqrt{2}}, du = \sqrt{2}dv$.

We know that $\int_{-\infty}^{\infty} \exp(-v^2)dv = \sqrt{\pi}$, due to a well-established result. Putting everything together, it follows that

$$\begin{aligned}\mathbb{E}(Y) &= \frac{1}{\sqrt{2\pi}} \cdot \sqrt{2e} \cdot \sqrt{\pi} \\ &= \sqrt{e}.\end{aligned}$$

Next, to find $\mathbb{V}(Y)$, we begin by computing $\mathbb{E}(Y^2)$. By the Law of the Unconscious Statistician and our previous computation of $f_Y(y)$, we obtain

$$\begin{aligned}\mathbb{E}(Y^2) &= \int_0^\infty y^2 \cdot \frac{1}{y\sqrt{2\pi}} \exp\left(-\frac{1}{2}\ln(y)^2\right) dy \\ &= \frac{1}{\sqrt{2\pi}} \int_0^\infty y \exp\left(-\frac{1}{2}\ln(y)^2\right) dy.\end{aligned}$$

Let $u = \ln y$, such that $y = e^u$ and $dy = e^u du$. Again, we deal with the integral separately from the constant factor, and multiply that back later. We have

$$\begin{aligned}\int_0^\infty y \exp\left(-\frac{1}{2}\ln(y)^2\right) dy &= \int_{-\infty}^\infty \exp\left(2u - \frac{1}{2}u^2\right) du \\ &= e^2 \int_{-\infty}^\infty \exp\left(-\left(\frac{1}{\sqrt{2}}u - \sqrt{2}\right)^2\right) du \\ &= e^2\sqrt{2} \int_{-\infty}^\infty \exp(-v^2)dv\end{aligned}$$

where like before, we substitute $v = \frac{1}{\sqrt{2}}u - \sqrt{2}$.

Again, using $\int_{-\infty}^{\infty} \exp(-v^2)dv = \sqrt{\pi}$, and putting everything together, we obtain

$$\begin{aligned}\mathbb{E}(Y^2) &= \frac{1}{\sqrt{2\pi}}e^2 \cdot \sqrt{2} \cdot \sqrt{\pi} \\ &= e^2.\end{aligned}$$

Thus

$$\begin{aligned}\mathbb{V}(Y) &= \mathbb{E}(Y^2) - \mathbb{E}(Y)^2 \\ &= e^2 - e\end{aligned}$$

and we are done. □

Problem 3.12. Prove the formulas given at the beginning of section 3.4, regarding mean and variance, for the following distributions: Bernoulli, Poisson, Uniform, Exponential, Gamma, Beta.

Solution. For each part of this solution, suppose X is a random variable distributed according to the distribution we are investigating.

For the **Bernoulli** distribution: we have $\mathbb{P}(X = 1) = p$ and $\mathbb{P}(X = 0) = 1 - p$. Thus $\mathbb{E}(X) = \sum xf(x) = 1 \cdot p + 0 \cdot (1 - p) = p$. Then, $\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$, and as $X^2 = X$ when $X = 0$ or $X = 1$, we obtain $\mathbb{V}(X) = \mathbb{E}(X) - \mathbb{E}(X^2) = p - p^2$.

For the **Poisson** distribution: we have $\mathbb{P}(X = x) = e^{-\lambda}\frac{\lambda^x}{x!}$. Thus we compute

$$\begin{aligned}\mathbb{E}(X) &= \sum_{x=0}^\infty x \cdot e^{-\lambda}\frac{\lambda^x}{x!} \\ &= e^{-\lambda} \cdot \lambda \cdot \sum_{x=1}^\infty \frac{\lambda^{x-1}}{(x-1)!} \\ &= e^{-\lambda} \cdot \lambda \cdot e^{\lambda} = \lambda.\end{aligned}$$

To find the variance, we first compute

$$\begin{aligned}
\mathbb{E}(X(X-1)) &= \sum_{x=0}^{\infty} x(x-1)e^{-\lambda}\frac{\lambda^x}{x!} \\
&= e^{-\lambda}\lambda^2 \sum_{x=2}^{\infty} \frac{\lambda^{x-2}}{(x-2)!} \\
&= \lambda^2.
\end{aligned}$$

Thus $\mathbb{E}(X^2 - X) = \lambda^2$, so using linearity of expectation and our previous work, we obtain $\mathbb{E}(X^2) = \lambda^2 + \lambda$. Finally, we have $\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$.

For the **Uniform** distribution: we have $f(x) = \frac{1}{b-a}$ for all $x \in (a, b)$. Thus

$$\begin{aligned}
\mathbb{E}(X) &= \int_a^b x \cdot \frac{1}{b-a} dx \\
&= \frac{1}{b-a} \int_a^b x dx \\
&= \frac{1}{b-a}\left(\frac{b^2}{2} - \frac{a^2}{2}\right) \\
&= \frac{a+b}{2}.
\end{aligned}$$

Then,

$$\begin{aligned}
\mathbb{E}(X^2) &= \int_a^b x^2 \cdot \frac{1}{b-a} dx \\
&= \frac{1}{b-a} \int_a^b x^2 dx \\
&= \frac{1}{b-a}\left(\frac{b^3}{3} - \frac{a^3}{3}\right) \\
&= \frac{1}{3}(a^2 + ab + b^2).
\end{aligned}$$

Thus

$$\begin{aligned}
\mathbb{V}(X) &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 \\
&= \frac{1}{3}(a^2 + ab + b^2) - \left(\frac{a+b}{2}\right)^2 \\
&= \frac{1}{12}(a^2 + b^2) - \frac{1}{6}ab \\
&= \frac{1}{12}(b-a)^2.
\end{aligned}$$

For the **Exponential** distribution: we have $f(x) = \frac{1}{\beta}e^{-x/\beta}$ when $x > 0$. Thus

$$\begin{aligned}
\mathbb{E}(X) &= \int_0^{\infty} x \cdot \frac{1}{\beta} e^{-x/\beta} dx \\
&= \frac{1}{\beta} \int_0^{\infty} x e^{-x/\beta} dx \\
&= \frac{1}{\beta}\left[\left[x \cdot (-\beta)\, e^{-x/\beta}\right]_0^{\infty} - \int_0^{\infty} (-\beta)\, e^{-x/\beta} dx\right] \\
&= \frac{1}{\beta}\left[\lim_{x\to\infty}\left(-\beta \cdot x e^{-x\beta}\right) + \beta^2 \int_0^{\infty} f(x) dx\right] \\
&= -\lim_{x\to\infty}\left(x e^{-x\beta}\right) + \beta
\end{aligned}$$

where we used integration by parts. Now we must simply compute this limit. Here we can use L'Hopital's rule:

$$\begin{aligned}\lim_{x\to\infty} xe^{-x\beta} &= \lim_{x\to\infty}\frac{x}{e^{x\beta}}\\ &= \lim_{x\to\infty}\frac{1}{\beta e^{x\beta}}\\ &= 0.\end{aligned}$$

Thus we obtain $\mathbb{E}(X) = \beta$. Next, we have

$$\begin{aligned}\mathbb{V}(X) &= \mathbb{E}(X^2) - \mathbb{E}(X)^2\\ &= \int_0^\infty x^2 \cdot \frac{1}{\beta}e^{-x/\beta}dx - \beta^2.\end{aligned}$$

We handle the integral separately, again with integration by parts:

$$\begin{aligned}\int_0^\infty x^2 e^{-x/\beta}dx &= \left[x^2(-\beta)e^{-x/\beta}\right]_0^\infty - \int_0^\infty (-\beta)e^{-x/\beta}\cdot 2xdx\\ &= -\beta\lim_{x\to\infty}\frac{x^2}{e^{x/\beta}} + 2\beta^2\int_0^\infty \frac{1}{\beta}e^{-x/\beta}\cdot xdx\\ &= 2\beta^2\cdot\mathbb{E}(X) = 2\beta^3\end{aligned}$$

where we use L'Hopital to take care of the limit and our work on $\mathbb{E}(X)$ to simplify the integral.

Thus

$$\begin{aligned}\mathbb{V}(X) &= \frac{1}{\beta}\cdot 2\beta^3 - \beta^2\\ &= \beta^2\end{aligned}$$

and we are done.

For the **Gamma** distribution: we have $f(x) = \frac{1}{\beta^\alpha\Gamma(\alpha)}x^{\alpha-1}e^{-x/\beta}$ for $x > 0$. Then

$$\begin{aligned}\mathbb{E}(X) &= \int_0^\infty x\cdot\frac{1}{\beta^\alpha\Gamma(\alpha)}x^{\alpha-1}e^{-x/\beta}dx\\ &= \int_0^\infty \frac{1}{\beta^\alpha\Gamma(\alpha)}x^\alpha e^{-x/\beta}dx\\ &= \beta\cdot\frac{\Gamma(\alpha+1)}{\Gamma(\alpha)}\int_0^\infty \frac{1}{\beta^{\alpha+1}\Gamma(\alpha+1)}x^\alpha e^{-x/\beta}dx\end{aligned}$$

where we use the fact that α, β are constants to move terms outside the integral. Note that what remains inside the integral is the PDF of a random variable distributed according to Gamma$(\alpha+1, \beta)$, so the integral itself evaluates to 1.

Thus, further simplifying the result requires evaluating $\frac{\Gamma(\alpha+1)}{\Gamma(\alpha)}$. We proceed now by integration by parts. Observe that

$$\begin{aligned}\Gamma(\alpha+1) &= \int_0^\infty y^\alpha e^{-y}dy\\ &= \left[y^\alpha(-e^{-y})\right]_0^\infty + \int_0^\infty e^{-y}y^{\alpha-1}\alpha dy\\ &= \lim_{y\to\infty}\left(-\frac{y^\alpha}{e^y}\right) + \alpha\int_0^\infty e^{-y}y^{\alpha-1}dy\\ &= \alpha\Gamma(\alpha)\end{aligned}$$

where we used L'Hopital to evaluate the limit. Thus $\mathbb{E}(X) = \alpha\beta$.

Evaluating the variance is similar. We have

$$\begin{aligned}\mathbb{E}(X^2) &= \int_0^\infty x^2 \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} dx \\ &= \beta^2 \cdot \frac{\Gamma(\alpha+2)}{\Gamma(\alpha)} \int_0^\infty \frac{1}{\beta^{\alpha+2}\Gamma(\alpha+2)} x^{\alpha+1} e^{-x/\beta} dx \\ &= \alpha(\alpha+1)\beta^2\end{aligned}$$

where we used our work previously to conclude that $\frac{\Gamma(\alpha+2)}{\Gamma(\alpha)} = \frac{\Gamma(\alpha+1)}{\Gamma(\alpha)} \cdot \frac{\Gamma(\alpha+2)}{\Gamma(\alpha+1)} = \alpha \cdot (\alpha+1)$.

Thus $\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \alpha(\alpha+1)\beta^2 - \alpha^2\beta^2 = \alpha\beta^2$.

Finally, for the **Beta** distribution: we have $f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}$.

Thus

$$\begin{aligned}\mathbb{E}(X) &= \int_0^1 x \cdot \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} dx \\ &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha+\beta+1)} \cdot \frac{\Gamma(\alpha+1)}{\Gamma(\alpha)} \int_0^1 \frac{\Gamma(\alpha+\beta+1)}{\Gamma(\alpha+1)\Gamma(\beta)} x^{\alpha}(1-x)^{\beta-1} dx \\ &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha+\beta+1)} \cdot \frac{\Gamma(\alpha+1)}{\Gamma(\alpha)} \\ &= \frac{\alpha}{\alpha+\beta}.\end{aligned}$$

using our previous work to simplify the result after eliminating the integral.

Also,

$$\begin{aligned}\mathbb{E}(X^2) &= \int_0^1 x^2 \cdot \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} dx \\ &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha+\beta+2)} \cdot \frac{\Gamma(\alpha+2)}{\Gamma(\alpha)} \int_0^1 \cdot \frac{\Gamma(\alpha+\beta+2)}{\Gamma(\alpha+2)\Gamma(\beta)} x^{\alpha+1}(1-x)^{\beta-1} dx \\ &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha+\beta+2)} \cdot \frac{\Gamma(\alpha+2)}{\Gamma(\alpha)} \\ &= \frac{1}{(\alpha+\beta)(\alpha+\beta+1)} \cdot \alpha \cdot (\alpha+1).\end{aligned}$$

Thus

$$\begin{aligned}\mathbb{V}(X) &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 \\ &= \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)} - \frac{\alpha^2}{(\alpha+\beta)^2} \\ &= \frac{\alpha(\alpha+1)(\alpha+\beta) - \alpha^2(\alpha+\beta+1)}{(\alpha+\beta)^2(\alpha+\beta+1)} \\ &= \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}.\end{aligned}$$

Having derived all the desired means and variances, we are finally done. □

Problem 3.13. Suppose we generate a random variable X in the following way. First, we flip a fair coin. If the coin is heads, take $X \sim \text{Uniform}(0,1)$. If the coin is tails, take $X \sim \text{Uniform}(3,4)$. Find the mean and standard deviation of X.

Solution. Let $Y \sim \text{Bernoulli}(\frac{1}{2})$ and $Z \sim \text{Uniform}(0,1)$ with Y and Z independent, such that $X = 3Y + Z$. It is easy to see this is an equivalent way of generating X.

Then

$$\begin{aligned}\mathbb{E}(X) &= \mathbb{E}(3Y+Z)\\ &= 3\mathbb{E}(Y)+\mathbb{E}(Z)\\ &= 3\cdot\frac{1}{2}+\frac{1}{2}\\ &= 2.\end{aligned}$$

Next,

$$\begin{aligned}\mathbb{V}(X) &= \mathbb{V}(3Y+Z)\\ &= 9\mathbb{V}(Y)+\mathbb{V}(Z)\\ &= 9\cdot\frac{1}{4}+\frac{1}{12}\\ &= \frac{7}{3}.\end{aligned}$$

Thus the standard deviation of X is $\sqrt{\frac{7}{3}}$. □

Problem 3.14. Let $X_1, \dots, X_m$ and $Y_1, \dots, Y_n$ be random variables and let $a_1, \dots, a_m$ and $b_1, \dots, b_n$ be constants. Show that

$$\mathrm{Cov}\left(\sum_{i=1}^m a_i X_i, \sum_{j=1}^n b_j Y_j\right) = \sum_{i=1}^m \sum_{j=1}^n a_i b_j \mathrm{Cov}(X_i, Y_j).$$

Solution. We have

$$\begin{aligned}\mathrm{Cov}\left(\sum_{i=1}^m a_i X_i, \sum_{j=1}^n b_j Y_j\right) &= \mathbb{E}\left(\sum_{i=1}^m a_i X_i \cdot \sum_{j=1}^n b_j Y_j\right) - \mathbb{E}\left(\sum_{i=1}^m a_i X_i\right)\mathbb{E}\left(\sum_{j=1}^n b_j Y_j\right)\\ &= \mathbb{E}\left(\sum_{i=1}^m \sum_{j=1}^n a_i b_i X_i Y_j\right) - \mathbb{E}\left(\sum_{i=1}^m a_i X_i\right)\mathbb{E}\left(\sum_{j=1}^n b_j Y_j\right)\\ &= \sum_{i=1}^m \sum_{j=1}^n a_i b_j \mathbb{E}(X_i Y_j) - \left(\sum_{i=1}^m a_i \mathbb{E}(X_i)\right)\left(\sum_{j=1}^n b_j \mathbb{E}(Y_j)\right).\\ &= \sum_{i=1}^m \sum_{j=1}^n \left[a_i b_j \mathbb{E}(X_i Y_j) - a_i \mathbb{E}(X_i) b_j \mathbb{E}(Y_j)\right].\end{aligned}$$

Note now that $a_i b_j \mathbb{E}(X_i Y_j) - a_i \mathbb{E}(X_i) b_j \mathbb{E}(Y_j) = a_i b_j \mathrm{Cov}(X_i, Y_j)$, so indeed,

$$\mathrm{Cov}\left(\sum_{i=1}^m a_i X_i, \sum_{j=1}^n b_j Y_j\right) = \sum_{i=1}^m \sum_{j=1}^n a_i b_j \mathrm{Cov}(X_i, Y_j).$$

□

Problem 3.15. Let

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{3}(x+y), & 0 \le x \le 1, 0 \le y \le 2\\ 0, & \text{otherwise.}\end{cases}$$

Find $\mathbb{V}(2X - 3Y + 8)$.

Solution. Begin by noting that $\mathbb{V}(2X-3Y+8) = \mathbb{E}(4X^2-12XY+9Y^2) - \mathbb{E}(2X-3Y)^2$, as the constant term is irrelevant.

Let us compute $f_X(x)$ and $f_Y(y)$. We have

$$\begin{aligned} f_X(x) &= \int f_{X,Y}(x,y)dy \\ &= \int_0^2 \frac{1}{3}(x+y)dy \\ &= \frac{2}{3}x + \frac{2}{3} \end{aligned}$$

and through similar methods we obtain $f_Y(y) = \frac{1}{3}y + \frac{1}{6}$.

Now, we will compute the expected values $\mathbb{E}(X), \mathbb{E}(X^2), \mathbb{E}(Y), \mathbb{E}(Y^2), \mathbb{E}(XY)$. That will suffice to obtain the desired variance.

$$\begin{aligned} \mathbb{E}(X) &= \int_0^1 x\left(\frac{2}{3}x + \frac{2}{3}\right)dx \\ &= \left[\frac{2}{9}x^3 + \frac{1}{3}x^2\right]_0^1 = \frac{5}{9}. \\ \mathbb{E}(X^2) &= \int_0^1 x^2\left(\frac{2}{3}x + \frac{2}{3}\right)dx \\ &= \left[\frac{1}{6}x^4 + \frac{2}{9}x^3\right]_0^1 = \frac{7}{18}. \\ \mathbb{E}(Y) &= \int_0^2 y\left(\frac{1}{3}y + \frac{1}{6}\right)dy \\ &= \left[\frac{1}{9}y^3 + \frac{1}{12}y^2\right]_0^2 = \frac{11}{9}. \\ \mathbb{E}(Y^2) &= \int_0^2 y^2\left(\frac{1}{3}y + \frac{1}{6}\right)dy \\ &= \left[\frac{1}{12}y^4 + \frac{1}{18}y^3\right]_0^2 = \frac{16}{9}. \\ \mathbb{E}(XY) &= \int\int xyf_{X,Y}(x,y)dydx \\ &= \int_0^1\int_0^2 \frac{1}{3}xy(x+y)dydx \\ &= \int_0^1\left(\frac{2}{3}x^2 + \frac{8}{9}x\right)dx = \frac{2}{3}. \end{aligned}$$

Thus

$$\begin{aligned} \mathbb{E}(4X^2-12XY+9Y^2) - \mathbb{E}(2X-3Y)^2 &= 4\mathbb{E}(X^2) - 12\mathbb{E}(XY) + 9\mathbb{E}(Y^2) - (2\mathbb{E}(X) - 3\mathbb{E}(Y))^2 \\ &= 4\cdot\frac{7}{18} - 12\cdot\frac{2}{3} + 9\cdot\frac{16}{9} - \left(2\cdot\frac{5}{9} - 3\cdot\frac{11}{9}\right)^2 \\ &= \frac{86}{9} - \frac{529}{81} = \frac{245}{81}. \end{aligned}$$

□

Problem 3.16. Let $r(x)$ be a function of x and let $s(y)$ be a function of y. Show that

$$\mathbb{E}(r(X)s(Y)|X) = r(X)\mathbb{E}(s(Y)|X).$$

Also, show that $\mathbb{E}(r(X)|X) = r(X)$.

Solution. We have $\mathbb{E}(r(X)s(Y)|X) = \int r(x)s(y)f_{Y|X}(y|x)dy$ and $r(X)\mathbb{E}(s(Y)|X) = r(X)\int s(y)f_{Y|X}(y|x)dy$. But as the integral is taken dy, the $r(x)$ term can be moved outside the integral in the first equation, so we are done.

As for the second part, take $s(Y) = 1$ in the equation we just proved. Then $\mathbb{E}(r(X)\cdot 1|X) = r(X)\mathbb{E}(1|X)$. But that reduces to $\mathbb{E}(r(X)|X) = r(X)$, as $\mathbb{E}(1|X) = 1$. $\square$

Problem 3.17. Prove that $\mathbb{V}(Y) = \mathbb{E}(\mathbb{V}(Y|X)) + \mathbb{V}(\mathbb{E}(Y|X))$.

Solution. Note that $\mathbb{V}(Y) = \mathbb{E}(Y^2) - \mathbb{E}(Y)^2$. Now we will use the iteration of expectation property.

We have

$$\begin{aligned}\mathbb{E}(Y^2) - \mathbb{E}(Y)^2 &= \mathbb{E}(\mathbb{E}(Y^2|X)) - \mathbb{E}(\mathbb{E}(Y|X))^2 \\ &= \mathbb{E}(\mathbb{V}(Y|X) - \mathbb{E}(Y|X)^2) - \mathbb{E}(\mathbb{E}(Y|X))^2\end{aligned}$$

where we use the fact that

$$\begin{aligned}\mathbb{V}(Y|X) &= \mathbb{E}((Y - \mathbb{E}(Y|X))^2|X) \\ &= \mathbb{E}(Y^2 - 2Y\mathbb{E}(Y|X) + \mathbb{E}(Y|X)^2|X) \\ &= \mathbb{E}(Y^2|X) - 2\mathbb{E}(Y\mathbb{E}(Y|X)) + \mathbb{E}(\mathbb{E}(Y|X)^2|X) \\ &= \mathbb{E}(Y^2|X) - 2\mathbb{E}(Y|X)^2 + \mathbb{E}(Y|X)^2 = \mathbb{E}(Y^2|X) - \mathbb{E}(Y|X)^2.\end{aligned}$$

(Note that the final line follows from our work in the previous problem: indeed, $2\mathbb{E}(Y\mathbb{E}(Y|X)) = 2\cdot\mathbb{E}(Y|X)\mathbb{E}(Y|X) = 2\mathbb{E}(Y|X)^2$, as $\mathbb{E}(r(X)s(Y)|X) = r(X)\mathbb{E}(s(Y)|X)$ and $\mathbb{E}(Y|X)$ is a function of X. Similarly, $\mathbb{E}(\mathbb{E}(Y|X)^2|X) = \mathbb{E}(Y|X)^2$, as $\mathbb{E}(r(X)|X) = r(X)$ and again, $\mathbb{E}(Y|X)^2$ is a function of X.)

Finally, observe that

$$\begin{aligned}\mathbb{E}(\mathbb{V}(Y|X) - \mathbb{E}(Y|X)^2) - \mathbb{E}(\mathbb{E}(Y|X))^2 &= \mathbb{E}(\mathbb{V}(Y|X)) - \mathbb{E}(\mathbb{E}(Y|X)^2) - \mathbb{E}(\mathbb{E}(Y|X))^2 \\ &= \mathbb{E}(\mathbb{V}(Y|X)) - \mathbb{V}(\mathbb{E}(Y|X))\end{aligned}$$

where we used linearity of expectation and the definition of $\mathbb{V}(\mathbb{E}(Y|X))$. $\square$

Problem 3.18. Show that if $\mathbb{E}(X|Y = y) = c$ for some constant c, then X and Y are uncorrelated.

Solution. Note that if $\mathbb{E}(X|Y = y) = c$ for some constant c, then the random variable $\mathbb{E}(X|Y)$ always takes on the value c.

Therefore, by iterating expectations, $\mathbb{E}(X) = \mathbb{E}(\mathbb{E}(X|Y)) = \mathbb{E}(c) = c$, and $\mathbb{E}(XY) = \mathbb{E}(\mathbb{E}(XY|Y)) = \mathbb{E}(Y\mathbb{E}(X|Y)) = \mathbb{E}(cY) = c\mathbb{E}(Y)$.

Thus we have

$$\mathrm{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = c\mathbb{E}(Y) - c\cdot\mathbb{E}(Y) = 0$$

and it follows that X and Y are uncorrelated. $\square$

Problem 3.20. Prove that if a is a vector and X is a random vector with mean μ and variance Σ, then $\mathbb{E}(a^T X) = a^T\mu$ and $\mathbb{V}(a^T X) = a^T\Sigma a$.

Prove that if A is a matrix, then $\mathbb{E}(AX) = A\mu$ and $\mathbb{V}(AX) = A\Sigma A^T$.

Solution. Note that if

$$X = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_k \end{pmatrix}$$

then its mean μ satisfies

$$\mu = \begin{pmatrix} \mathbb{E}(X_1) \\ \mathbb{E}(X_2) \\ \vdots \\ \mathbb{E}(X_k) \end{pmatrix}$$

and its variance-covariance matrix Σ satisfies

$$\Sigma = \begin{pmatrix} \mathbb{V}(X_1) & \mathrm{Cov}(X_1, X_2) & \cdots & \mathrm{Cov}(X_1, X_k) \\ \mathrm{Cov}(X_2, X_1) & \mathbb{V}(X_2) & \cdots & \mathrm{Cov}(X_2, X_k) \\ \vdots & \vdots & \vdots & \vdots \\ \mathrm{Cov}(X_k, X_1) & \mathrm{Cov}(X_k, X_2) & \cdots & \mathbb{V}(X_k) \end{pmatrix}.$$

Suppose that $a = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_k \end{pmatrix}$. Then

$$\begin{aligned} a^T X &= \begin{pmatrix} a_1 & a_2 & \cdots & a_k \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_k \end{pmatrix} \\ &= a_1 X_1 + a_2 X_2 + \ldots + a_k X_k \end{aligned}$$

so therefore by linearity of expectation,

$$\begin{aligned} \mathbb{E}(a^T X) &= \mathbb{E}(a_1 X_1 + \ldots + a_k X_k) \\ &= \mathbb{E}(a_1 X_1) + \ldots + \mathbb{E}(a_k X_k) \\ &= a_1 \mathbb{E}(X_1) + \ldots + a_k \mathbb{E}(X_k) \\ &= a^T \mu. \end{aligned}$$

Now, note that

$$\mathbb{V}(a^T X) = \mathbb{V}(a_1 X_1 + \ldots + a_j X_j)$$

and that

$$\begin{aligned} a^T \Sigma a &= \begin{pmatrix} a_1 & a_2 & \cdots & a_k \end{pmatrix} \begin{pmatrix} \mathbb{V}(X_1) & \mathrm{Cov}(X_1, X_2) & \cdots & \mathrm{Cov}(X_1, X_k) \\ \mathrm{Cov}(X_2, X_1) & \mathbb{V}(X_2) & \cdots & \mathrm{Cov}(X_2, X_k) \\ \vdots & \vdots & \vdots & \vdots \\ \mathrm{Cov}(X_k, X_1) & \mathrm{Cov}(X_k, X_2) & \cdots & \mathbb{V}(X_k) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_k \end{pmatrix} \\ &= \begin{pmatrix} \sum_{i=1}^k a_i \mathrm{Cov}(X_i, X_1) & \cdots & \sum_{i=1}^k a_i \mathrm{Cov}(X_i, X_k) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_k \end{pmatrix} \\ &= \sum_{j=1}^k \sum_{i=1}^k a_i a_j \mathrm{Cov}(X_i, X_j) = \sum_{i=1}^k a_i^2 \mathbb{V}(X_i) + 2 \sum_{1 \le i < j \le k} a_i a_j \mathrm{Cov}(X_i, X_j). \end{aligned}$$

But that means that $\mathbb{V}(A^T X) = a^T \Sigma a$, by Theorem 3.20.

For the second part of the problem, suppose that $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{pmatrix}$. Then

$$AX = \begin{pmatrix} \sum_{i=1}^k a_{1i} X_i \\ \sum_{i=1}^k a_{2i} X_i \\ \vdots \\ \sum_{i=1}^k a_{ki} X_i \end{pmatrix}$$

so thus

$$
\begin{aligned}
\mathbb{E}(AX) &= \begin{pmatrix} \mathbb{E}\left(\sum_{i=1}^{k} a_{1i}X_i\right) \\ \mathbb{E}\left(\sum_{i=1}^{k} a_{2i}X_i\right) \\ \vdots \\ \mathbb{E}\left(\sum_{i=1}^{k} a_{ki}X_i\right) \end{pmatrix} \\
&= \begin{pmatrix} \sum_{i=1}^{k} a_{1i}\mathbb{E}(X_i) \\ \sum_{i=1}^{k} a_{2i}\mathbb{E}(X_i) \\ \vdots \\ \sum_{i=1}^{k} a_{ki}\mathbb{E}(X_i) \end{pmatrix} \\
&= A\mu
\end{aligned}
$$

by linearity of expectation.

To show that $\mathbb{V}(AX) = A\Sigma A^T$, we note first that

$$
\begin{aligned}
(X - \mathbb{E}(X))(X - \mathbb{E}(X))^T &= \begin{pmatrix} X_1 - \mathbb{E}(X_1) \\ \vdots \\ X_k - \mathbb{E}(X_k) \end{pmatrix} \begin{pmatrix} X_1 - \mathbb{E}(X_1) & \cdots & X_k - \mathbb{E}(X_k) \end{pmatrix} \\
&= \begin{pmatrix} (X_1 - \mathbb{E}(X_1))(X_1 - \mathbb{E}(X_1)) & \cdots & (X_1 - \mathbb{E}(X_1))(X_k - \mathbb{E}(X_k)) \\ \vdots & \vdots & \vdots \\ (X_k - \mathbb{E}(X_k))(X_1 - \mathbb{E}(X_1)) & \cdots & (X_k - \mathbb{E}(X_k))(X_k - \mathbb{E}(X_k)) \end{pmatrix}
\end{aligned}
$$

which shows that, as $\mathrm{Cov}(A, B) = \mathbb{E}[(A - \mathbb{E}(A))(B - \mathbb{E}(B))]$ for random variables A and B, we have $\mathbb{E}[(X - \mathbb{E}(X))(X - \mathbb{E}(X))^T] = \mathbb{V}(X)$.

Thus

$$
\begin{aligned}
\mathbb{V}(AX) &= \mathbb{E}\left[(AX - \mathbb{E}(AX))(AX - \mathbb{E}(AX))^T\right] \\
&= \mathbb{E}\left[(AX - A\mathbb{E}(X))(AX - A\mathbb{E}(X))^T\right] \\
&= \mathbb{E}\left[A(X - \mathbb{E}(X))(X - \mathbb{E}(X))^T A^T\right] \\
&= A\mathbb{V}(X)A^T
\end{aligned}
$$

where we used the properties $A\mathbb{E}(X) = \mathbb{E}(AX)$ (the previous result) and $(AB)^T = B^T A^T$. □

Problem 3.21. Let X and Y be random variables. Suppose that $\mathbb{E}(Y|X) = X$. Show that $\mathrm{Cov}(X, Y) = \mathbb{V}(X)$.

Solution. As $\mathbb{E}(Y|X) = X$, we know that $Y = \mathbb{E}(\mathbb{E}(Y|X)) = \mathbb{E}(X)$.

That means that

$$
\begin{aligned}
\mathrm{Cov}(X, Y) &= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) \\
&= \mathbb{E}(\mathbb{E}(XY|X)) - \mathbb{E}(X)\mathbb{E}(\mathbb{E}(X)) \\
&= \mathbb{E}(X\mathbb{E}(Y|X)) - \mathbb{E}(X)^2 \\
&= \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \mathbb{V}(X)
\end{aligned}
$$

as desired. Note that we used the fact that $\mathbb{E}(\mathbb{E}(X)) = \mathbb{E}(X)$; this follows by taking Z to be a random variable independent of X, and then noting that $\mathbb{E}(X) = \mathbb{E}(\mathbb{E}(X|Z)) = \mathbb{E}(\mathbb{E}(X))$ through the law of iterated expectations and the fact that $\mathbb{E}(X|Z) = \mathbb{E}(X)$. □

Problem 3.22. Let $X \sim \text{Uniform}(0,1)$. Let $0 < a < b < 1$. Let

$$Y = \begin{cases} 1, & 0 < x < b \\ 0 & \text{otherwise} \end{cases}$$

and let

$$Z = \begin{cases} 1, & a < x < 1 \\ 0, & \text{otherwise.} \end{cases}$$

Are Y and Z independent? Also, find $\mathbb{E}(Y|Z)$.

Solution. We have $\mathbb{P}(Y=1) = b$ and $\mathbb{P}(Z=1) = 1-a$. But $\mathbb{P}(Y=1, Z=1) = \mathbb{P}(a < X < b) = b-a$. Thus $\mathbb{P}(Y=1)\mathbb{P}(Z=1) = b(1-a) \neq b-a = \mathbb{P}(Y=1, Z=1)$, so Y and Z are not independent.

Note that we have $\mathbb{P}(Y=1|Z=1) = \mathbb{P}(0 < x < b | a < x < 1) = \frac{b-a}{1-a}$, and $\mathbb{P}(Y=0|Z=1) = 1 - \frac{b-a}{1-a} = \frac{1-b}{1-a}$. Thus $\mathbb{E}(Y|Z=1) = 1 \cdot \frac{b-a}{1-a} + 0 \cdot \frac{1-b}{1-a} = \frac{b-a}{1-a}$.

Moreover, we have $\mathbb{P}(Y=1|Z=0) = 1$, so $\mathbb{E}(Y|Z=0) = 1$.

Thus it follows that

$$\mathbb{E}(Y|Z) = \begin{cases} 1, & Z = 0 \\ \frac{b-a}{1-a}, & Z = 1. \end{cases}$$

□

Problem 3.23. Find the moment generating function for the Poisson, Normal, and Gamma distributions.

Solution. For each part of this solution, suppose X is a random variable distributed according to the distribution we are investigating.

For the **Poisson** distribution: we have $\mathbb{P}(X = x) = e^{-\lambda}\frac{\lambda^x}{x!}$. Thus we wish to compute $\psi_X(t) = \sum_{x=0}^{\infty} e^{tx} e^{-\lambda}\frac{\lambda^x}{x!}$. The key observation is that $e^{tx} = (e^t)^x$, and with this in mind, we see that

$$\begin{aligned}
\phi_X(t) = \sum_{x=0}^{\infty} e^{tx} e^{-\lambda}\frac{\lambda^x}{x!} &= e^{-\lambda}\sum_{x=0}^{\infty}\frac{(e^t)^x\lambda^x}{x!} \\
&= e^{-\lambda}\sum_{x=0}^{\infty}\frac{(e^t\lambda)^x}{x!} \\
&= e^{-\lambda}e^{e^t\lambda} \\
&= e^{\lambda(e^t-1)}.
\end{aligned}$$

For the **Normal** distribution: we have $f_X(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right)$. Thus we wish to compute $\psi_X(t) = \int_{-\infty}^{\infty} e^{tx}\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right)dx$.

We have

$$\begin{aligned}
\int_{-\infty}^{\infty}\exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2 + tx\right)dx &= \int_{-\infty}^{\infty}\exp\left(-\frac{1}{2\sigma^2}x^2 + \left(\frac{\mu}{\sigma^2}+t\right)x - \frac{\mu^2}{2\sigma^2}\right)dx \\
&= \int_{-\infty}^{\infty}\exp\left(-\left(\frac{1}{\sigma\sqrt{2}}x - \frac{(\frac{\mu}{\sigma^2}+t)\sigma\sqrt{2}}{2}\right)^2 + \left(t\mu + \frac{t^2\sigma^2}{2}\right)\right)dx \\
&= e^{\mu t + \frac{t^2\sigma^2}{2}}\int_{-\infty}^{\infty}\exp\left(-\left(\frac{1}{\sigma\sqrt{2}}x - \frac{(\frac{\mu}{\sigma^2}+t)\sigma\sqrt{2}}{2}\right)^2\right)dx \\
&= e^{\mu t + \frac{t^2\sigma^2}{2}}\cdot\sigma\sqrt{2}\int_{-\infty}^{\infty}\exp(-u^2)du = e^{\mu t + \frac{t^2\sigma^2}{2}}\cdot\sigma\sqrt{2\pi}
\end{aligned}$$

where we completed the square and u-substituted.

Thus we obtain $\phi_X(t) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{\mu t + \frac{t^2\sigma^2}{2}} \cdot \sigma\sqrt{2\pi} = e^{\mu t + \frac{t^2\sigma^2}{2}}$.

For the **Gamma** distribution: we have $f_X(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}$. Thus we wish to compute $\psi_X(t) = \int_0^\infty e^{tx} \frac{1}{\beta^\alpha\Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} dx$. In attempting to evaluate the integral, we will make use of the substitution $-x/\beta + tx = -x/k$, or equivalently, $k = \frac{1}{\frac{1}{\beta} - t}$ or $\frac{k}{1+tk} = \beta$.

We have

$$\begin{aligned}
\int_0^\infty e^{tx} \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} dx &= \int_0^\infty \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta + tx} \\
&= \int_0^\infty \frac{1}{\left(\frac{k}{1+tk}\right)^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/k} dx \\
&= (1+tk)^\alpha
\end{aligned}$$

where we used that the integral, after extracting $(1+tk)^\alpha$, is just the integral of a Gamma(α, k) distribution. (Note that this is valid only when $\frac{1}{\beta} > t$, otherwise k is negative.)

Then, substituting back in for k yields

$$\begin{aligned}
\psi_X(t) = (1+tk)^\alpha &= \left(1 + \frac{t}{\frac{1}{\beta} - t}\right)^\alpha \\
&= \left(\frac{1/\beta}{1/\beta - t}\right)^\alpha \\
&= \left(\frac{1}{1-\beta t}\right)^\alpha
\end{aligned}$$

and we are done. □

Problem 3.24. Let $X_1, \ldots, X_n \sim \text{Exp}(\beta)$. Find the moment generating function of X_i. Prove that $\sum_{i=1}^n X_i \sim \text{Gamma}(n, \beta)$.

Solution. Note that the exponential distribution is a special case of the Gamma distribution. Thus, from our work on the previous problem we know that $\psi_{X_i}(t) = \frac{1}{1-\beta t}$.

Now, we know that $Y = \sum_{i=1}^n X_i$ satisfies $\psi_Y(t) = \prod_{i=1}^n \psi_{X_i}(t) = \left(\frac{1}{1-\beta t}\right)^n$, which is the moment generating function of the Gamma(n, β) distribution. Therefore $Y \sim \text{Gamma}(n, \beta)$. □