Probability and Statistics Homework #12

Ari Feiglin

Question 12.1:

Let X be the number of corn kernels which didn't become popcorn in a certain bag. Suppose $X \sim \mathcal{N}(25, 25)$.

(1) What is the average number of kernels which didn't pop?

(2) What is the probability that in a single bag of kernels, more than 30 kernels didn't pop?

(3) What is the average number of kernels in the 3rd decile which didn't pop?

(1) Since $X \sim \mathcal{N}(25, 25)$, $\mathbb{E}[X] = 25$.

(2) We want to compute:

$$\mathbb{P}(X > 30) = 1 - \mathbb{P}(X \leq 30) = 1 - F_X(30)$$

And since $X \sim \mathcal{N}(25, 25)$, $F_X(t) = \Phi\left(\frac{t - \mu_X}{\sigma_X}\right)$. So:

$$1 - F_X(30) = 1 - \Phi\left(\frac{30 - 25}{5}\right) = 1 - \Phi(1)$$

And using the Z-Table, we see that $\Phi(1) \approx 0.84134$, so:

$$\mathbb{P}(X > 30) \approx 0.1587$$

(3) The third decile is the event when $\mathbb{P}(X \leq t) \geq 0.3$. Let $Y \sim \mathcal{N}(25, 25)$. This means that $X \overset{d}{=} 5Y + 25$, so $\mathbb{P}(X \leq t) \geq 0.3$ if and only if:

$$\mathbb{P}(5Y + 25 \leq t) \geq 0.3 \iff \mathbb{P}\left(Y \leq \frac{t}{5} - 5\right) \geq 0.3 \iff \Phi\left(\frac{t}{5} - 5\right) \geq 0.3$$

Let $s = \frac{t}{5} - 5$ for when both sides are equal, so something is in the third percentile if and only if $X \geq t$. The two sides equal approximately when $s \approx -0.52$, so when $t \approx 22.4$. So we want to find the average value for $t \geq 22.4$. This is equal to:

$$\int_t^\infty x f_X(x)\, dx = \frac{1}{5\sqrt{2\pi}} \int_t^\infty x e^{-\frac{(x-25)^2}{50}}\, dx$$

Substituting $u = (x - 25)^2$ into the integral yields:

$$\frac{1}{10\sqrt{2\pi}} \int_{(t-25)^2}^\infty e^{-u/50}\, du + 25 \frac{1}{5\sqrt{2\pi}} \int_t^\infty e^{-\frac{(x-25)^2}{50}}\, dx$$

The right side is equal to $25\,\mathbb{P}(X \geq t) = 25(1 - F_X(t)) = 25\left(1 - \Phi\left(\frac{t-25}{5}\right)\right)$. Plugging in $t = 22.4$ gives 17.46175. The left side is equal to:

$$\frac{50}{10\sqrt{2\pi}} e^{-\frac{(t-25)^2}{50}} \approx 1.74$$

So all in all this is equal to about 19.2.

Question 12.2:

X and Y are two independent random variables which have a distribution of $\text{Exp}(1)$. Show that $\min\{X, Y\} \sim \text{Exp}(2)$.

Let's prove something more general: if X and Y are two independent random variables and $X, Y \sim \text{Exp}(\lambda)$, then $\min\{X, Y\} \sim \text{Exp}(2\lambda)$. If this is true, then setting $\lambda = 1$ proves the desired statement.
Let:

$$Z \coloneqq \min\{X, Y\}$$

We will compute $\overline{F}_Z$:

$$\overline{F}_Z(t) = \mathbb{P}(Z \geq t) = \mathbb{P}(\min\{X, Y\} \geq t) = \mathbb{P}(X \geq t, Y \geq t)$$

And since X and Y are independent, this is equal to:

$$= \mathbb{P}(X \geq t) \cdot \mathbb{P}(Y \geq t) = \overline{F}_X(t) \cdot \overline{F}_Y(t)$$

Since $F_X(t) = 1 - e^{-\lambda t}$, this means that $\overline{F}_X(t) = e^{-\lambda t}$. And since both X and Y have the same distribution, $\overline{F}_X = \overline{F}_Y$. So:

$$\overline{F}_Z(t) = e^{-\lambda t} \cdot e^{-\lambda t} = e^{-2\lambda t}$$

This means that:

$$F_Z(t) = 1 - \overline{F}_Z(t) = 1 - e^{-2\lambda t}$$

And so:

$$f_Z(t) = F_Z'(t) = 2\lambda e^{-2\lambda t}$$

Which is the probability density function of $\text{Exp}(2\lambda)$, so $Z \sim \text{Exp}(2\lambda)$, as required.

Question 12.3:

Suppose X and Y are independent random variables such that $X, Y \sim \operatorname{Exp}(1)$. Show that $X+Y$ has a probability densty function of te^{-t}.

Suppose instead that $X, Y \sim \operatorname{Exp}(\lambda)$. And let $Z \coloneqq Z + Y$. Then we know by the convolution theorem that:

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(t) \cdot f_Y(z-t)\, dt$$

Since X and Y distribute geometrically, $f_X(t) = 0$ if $t < 0$ and $f_Y(z-t) = 0$ if $z - t < 0 \iff t > z$. So the inetgrand is zero outside the range $0 \leq t \leq z$. So this integral is equal to:

$$= \int_0^z f_X(t) \cdot f_Y(z-t)\, dt = \int_0^z \lambda e^{-\lambda t} \cdot \lambda e^{-\lambda(z-t)}\, dt = \lambda^2 \int_0^z e^{-\lambda z}\, dt = z\lambda^2 e^{-\lambda z}$$

So if $\lambda = 1$ (as it is in the question):

$$f_Z(z) = ze^{-z}$$

As required.

Question 12.4:

Suppose $\theta \sim \mathrm{Unif}(0, 2\pi)$. Let $X = \cos\theta$ and $Y = \sin\theta$.

(1) Show that X and Y are uncorrelated.

(2) Show that X and Y are dependent.

(1) Notice that $X \cdot Y = \cos\theta \cdot \sin\theta = \frac{1}{2}\sin(2\theta)$. So:

$$\mathbb{E}[XY] = \frac{1}{2}\int_0^{2\pi} \sin(2x)\, f_\theta(x)\, dx = \frac{1}{2}\int_0^{2\pi} \sin(2x)\, \frac{1}{2\pi}\, dx = \frac{1}{4\pi}\int_0^{2\pi} \sin(2x)\ dx$$

And:

$$\int_0^{2\pi} \sin(2x)\ dx = -\frac{1}{2}\cos(2x)\Big|_0^{2\pi} = 0$$

So $\mathbb{E}[XY] = 0$.
And:

$$\mathbb{E}[Y] = \int_0^{2\pi} \sin(x) \cdot \frac{1}{2\pi}\, dx = -\frac{1}{2\pi}\cos(x)\Big|_0^{2\pi} = 0$$

And so:

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

So X and Y are not correlated, as required.

(2) Let's take a look at the probability that $\mathbb{P}(X, Y \leq 0.5)$. We can compute that $X \leq 0.5$ when $0 \leq \theta\frac{\pi}{6}$ or $\frac{5\pi}{6} \leq \theta \leq 2\pi$. And that $Y \leq 0.5$ when $\frac{\pi}{3} \leq \theta \leq \frac{5\pi}{3}$.
And since θ is uniform (so we can compute the lengths of the intervals where the event occurs and divide by 2π):

$$\mathbb{P}(X \leq 0.5) = \frac{\frac{\pi}{6} + 2\pi - \frac{5\pi}{6}}{2\pi} = \frac{2}{3}$$

And:

$$\mathbb{P}(Y \leq 0.5) = \frac{\frac{5\pi}{3} - \frac{\pi}{3}}{2\pi} = \frac{2}{3}$$

But $X, Y \leq 0.5$ occurs only if $\frac{5\pi}{6} \leq \theta \leq \frac{5\pi}{3}$, so:

$$\mathbb{P}(X, Y \leq 0.5) = \frac{\frac{5\pi}{3} - \frac{5\pi}{6}}{2\pi} = \frac{5}{12} \neq \frac{4}{9} = \mathbb{P}(X \leq 0.5) \cdot \mathbb{P}(Y \leq 0.5)$$

So X and Y are dependent.

Question 12.5:

Compute the moment-generating function of the following distributions:

(1) $X \sim \text{Unif}\,[1, \dots, n]$

(2) $Y \sim \text{Unif}\,[a, b]$

(3) $Z \sim \text{Geo}\,(p)$

(4) $W \sim \text{Exp}\,(\lambda)$

(1) By using the definition of the moment-generating function and LOTUS:

$$M_X(t) = \mathbb{E}\left[e^{tX}\right] = \sum_{k=1}^{n} e^{tk} \cdot \frac{1}{n} = \frac{1}{n}\sum_{k=1}^{n} \left(e^t\right)^k$$

This is a geometric sum, which is equal to:

$$\frac{1}{n} \cdot \frac{e^t(e^{tn} - 1)}{e^t - 1}$$

So:

$$M_X(t) = \frac{1}{n} \cdot \frac{e^{(n+1)t} - e^t}{e^t - 1}$$

(2) Again, by definition:

$$M_Y(t) = \mathbb{E}\left[e^{tY}\right] = \int_a^b e^{tx} \cdot \frac{1}{b-a}\, dx = \frac{1}{b-a}\frac{e^{tx}}{t}\bigg|_a^b$$

So:

$$M_Y(t) = \frac{e^{tb} - e^{ta}}{t(b-a)}$$

(3) Not entirely sure what to write here, it seems kind of repetitive.

$$M_Z(t) = \mathbb{E}\left[e^{tZ}\right] = \sum_{k=1}^{\infty} e^{tk} \cdot p(1-p)^{k-1} = pe^t \sum_{k=1}^{\infty} \left(e^t(1-p)\right)^{k-1} = pe^t \sum_{k=0}^{\infty} \left(e^t(1-p)\right)^k$$

This is an infinite geometric sum, which converges only if $e^t(1-p) < 1 \iff e^{-t} > 1-p \iff t < -\log(1-p)$

So if $t < -\log(1-p)$, then:

$$M_Z(t) = \frac{pe^t}{1-(1-p)e^t}$$

(4) ╲○╱

$$M_W(t) = \mathbb{E}\left[e^{tW}\right] = \int_0^{\infty} e^{tx}\lambda e^{-\lambda x}\, dx = \lambda \int_0^{\infty} e^{x(t-\lambda)}\, dx$$

If $t = \lambda$, this is equal to the integral of 1, which diverges. Otherwise, it equals to:

$$= \lambda \frac{e^{x(t-\lambda)}}{t-\lambda}\bigg|_0^{\infty}$$

This converges only if $t - \lambda < 0$, so if $t < \lambda$. If so:

$$= \lambda \cdot -\frac{1}{t-\lambda} = \frac{\lambda}{\lambda - t}$$

So for $t < \lambda$:

$$M_W(t) = \frac{\lambda}{\lambda - t}$$

Question 12.6:

Suppose Z is a random variable with the following density function:

$$f_Z(x) = \begin{cases} xe^{-x} & x > 0 \\ 0 & \text{else} \end{cases}$$

And suppose U is another random variable such that $U \mid Z \sim \text{Unif}\,[0, Z]$. Prove that $U \sim \text{Exp}\,(1)$.

From the law of total probability, we know that:

$$f_U(t) = \int_{-\infty}^{\infty} f_{U|Z=z}(t) \cdot f_Z(z)\, dz$$

And since $U \mid Z \sim \text{Unif}\,[0, Z]$, this means that $t \leq z$ and $z \geq 0$ (since $f_Z(z)$ is 0 for $z < 0$). And if $t < 0$, then $f_{U|Z=z}(t) = 0$, since it is uniform in $[0, z]$. So for $t \geq 0$, this is equal to:

$$f_U(t) = \int_t^{\infty} \frac{1}{z} \cdot ze^{-z}\, dz = \int_t^{\infty} e^{-z}\, dz = -e^{-z}\Big|_t^{\infty} = e^{-t}$$

Which is the probability density function for $\text{Exp}\,(1)$, so $U \sim \text{Exp}\,(1)$, as required.

Question 12.7:

Suppose (X, Y, Z) is a triplet of random variables such that:

$$f_{X,Y,Z}(x,y,z) = \begin{cases} ke^{-(ax+by+cz)} & x,y,z > 0 \\ 0 & \text{else} \end{cases}$$

(1) What is the value of k?

(2) Find the joint probability function of X and Y.

(3) Find the probability density function of X.

(4) Are X, Y, and Z independent?

$a, b, c > 0$ are constants.

(1) We know that:

$$\iiint_{\mathbb{R}^3} f_{X,Y,Z}(x,y,z)\,dx\,dy\,dz = 1$$

So:

$$1 = \int_0^\infty \int_0^\infty \int_0^\infty ke^{-(ax+by+cz)}\,dx\,dy\,dz = k\int_0^\infty e^{-cz} \int_0^\infty e^{-by} \int_0^\infty e^{-ax}\,dx\,dy\,dz =$$

$$= k\int_0^\infty e^{-cz}\,dz \int_0^\infty e^{-cy}\,dy \int_0^\infty e^{-ax}\,dx$$

Now, notice that:

$$\int_0^\infty e^{-ax}\,dx = -\frac{e^{-ax}}{a}\bigg|_0^\infty = \frac{1}{a}$$

So:

$$1 = \frac{k}{abc}$$

Multiplying both sides by abc yields:

$$k = abc$$

(2) We know that:

$$f_{X,Y}(x,y) = \int_{-\infty}^\infty f_{X,Y,Z}(x,y,z)\,dz$$

So if $x < 0$ or $y < 0$, this is equal to 0, since $f_{X,Y,Z}(x,y,z) = 0$ in this case. Otherwise:

$$= k\int_0^\infty e^{-(ax+by+cz)}\,dz = ke^{-(ax+by)}\int_0^\infty e^{-cz}\,dz = e^{-(ax+by)}\cdot\frac{k}{c} = abe^{-(ax+by)}$$

So:

$$f_{X,Y}(x,y) = \begin{cases} ab\cdot e^{-(ax+by)} & x,y > 0 \\ 0 & \text{else} \end{cases}$$

(3) Similar to the previous subquestion, we know that if $x < 0$, $f_X(x) = 0$. And otherwise:

$$f_X(x) = \int_{-\infty}^\infty f_{X,Y}(x,y)\,dy = ab\int_0^\infty e^{-(ax+by)}\,dy = ab\cdot e^{-ax}\int_0^\infty e^{-by}\,dy = ae^{-ax}$$

So:

$$X = \begin{cases} ae^{-ax} & x > 0 \\ 0 & \text{else} \end{cases}$$

So in other words, $X \sim \text{Exp}\,(a)$.

(4) Similar to the process done in the previous question, we can see that $Y \sim \text{Exp}(b)$ and $Z \sim \text{Exp}(c)$. (This can also be proven by symmetry/no loss of generality). And $f_{X,Z}(x,z) = ace^{-(ax+cz)}$ and $f_{Y,Z}(y,z) = bce^{-(by+cz)}$. This means that $f_{X,Y} = f_X \cdot f_Y$, $f_{Y,Z} = f_Y \cdot f_Z$, $f_{X,Z} = f_X \cdot f_Z$, and $f_{X,Y,Z} = f_X \cdot f_Y \cdot f_Z$. So X, Y, and Z are independent.

Question 12.8:

Prove the AM-GM inequality:

$$(a_1 \cdot a_2 \cdots a_n)^{\frac{1}{n}} \leq \frac{a_1 + a_2 + \cdots + a_n}{n}$$

Let $X \sim \text{Unif}\{a_1, \ldots, a_n\}$. Then we know that $\mathbb{E}[X] = \frac{a_1 + \cdots + a_n}{n}$. Given a function f, we know that

$$\mathbb{E}[f(X)] = \frac{f(a_1) + \cdots + f(a_n)}{n}$$

So we want a convex function which will somehow get us a product of the a_is. We can use $f(x) = -\log x$. This is convex since its second derivative is strictly positive. So:

$$\mathbb{E}[f(X)] = -\frac{\log(a_1) + \cdots + \log(a_n)}{n} = -\frac{1}{n}\log(a_1 \cdots a_n) = -\log\left((a_1 \cdots a_n)^{\frac{1}{n}}\right)$$

And:

$$f(f)\,\mathbb{E}[X] = -\log\left(\frac{a_1 + \cdots + a_n}{n}\right)$$

By Jensen's inequality, since f is convex:

$$f(f)\,\mathbb{E}[X] \leq \mathbb{E}[f(X)]$$

So:

$$-\log\left(\frac{a_1 + \cdots + a_n}{n}\right) \leq -\log\left((a_1 \cdots a_n)^{\frac{1}{n}}\right)$$

Multiplying both sides by -1 and exponentiating yields:

$$\frac{a_1 + \cdots + a_n}{n} \geq (a_1 \cdots a_n)^{\frac{1}{n}}$$

As required.