Mathematical Statistics
Problem Set #2

Max Thrush Hukill

February 7, 2019

Collaborators: Sam Harder (LaTeX tutor), Connor Fitch, Juliana Taube, Gavin Shilling, Dr. Jack

1. **Rounded joints.** Define a joint density function as $f_{X,Y}(x,y) = c$ if $x^2 + y^2 \leq R^2$ and zero outside of that region.

   (a) Find the value of c that ensures that this function is a probability density function (PDF).

   (b) Find the marginal density of X and Y.

   (c) Suppose D is a distance from the origin. Find $\mathbb{P}(D \leq a)$. Note that $D = \sqrt{X^2 + Y^2}$.

   (d) What is $\mathbb{E}(D)$?

Solution. In order for $f_X(x)$ to be a probability density function, the area under the distribution curve must equal one. By inspection, we see that the join density function graphs a cylinder of height c over a circle in the xy-plane with radius R. Indeed, we know the volume for a cylinder can be written as the base times height, where the base is the area of a circle of radius R. Thus, we have

$$\begin{aligned} \text{Base} \cdot \text{Height} &= \text{Volume} \\ (\pi R^2)(c) &= 1 \\ &\boxed{c = \frac{1}{\pi R^2}.} \end{aligned}$$

To find the marginal density of X and Y, simply integrate the joint density function with respect to each variable like so,

$$\begin{aligned} f_Y(y) &= \int_X f_{X,Y}(x,y)\ dx \\ &= \int_{-\sqrt{R^2-y^2}}^{\sqrt{R^2-y^2}} c\ dx = 2c\sqrt{R^2-x^2} = \boxed{\frac{2\sqrt{R^2-x^2}}{\pi R^2}.} \end{aligned}$$

Analogously,

$$f_X(x) = \int_Y f_{X,Y}(x,y)dy$$
$$= \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} cdy = 2c\sqrt{R^2-y^2} = \boxed{\frac{2\sqrt{R^2-y^2}}{\pi R^2}}.$$

Now consider D, a new random variable corresponding to a distance from the origin. We want to find $\mathbb{P}(D \leq a)$. There are two cases to consider. The first case is trivial; when $a > R$, the $\mathbb{P}(D \leq a) = 1$. This simply follows from the fact that D cannot exceed R by construction (recall the note, $D = \sqrt{X^2+Y^2} \leq R$). But when $a \leq R$, we ought to follow our instincts for geometry. Imagine this case as a series of concentric cylinders centered at the origin, wherein the outermost cylinder has radius R, and there is some other set cylinder with radius a (recall, the case assumes $a \leq R$). The third cylinder has radius D. In effect, we want to find the cumulative distribution function (CDF) of all D that satisfy $D \leq a$, which will tell us the $\mathbb{P}(D \leq a)$. Again, employ the formula of a cylinder, with constant height c, to see

$$CDF = \mathbb{P}(D \leq a) = c\pi a^2 = \frac{1}{\pi R^2}\pi a^2 = \boxed{\frac{a^2}{R^2}}$$

Finally, to find the expectation of D, we differentiate the CDF with respect to a to recover the PDF, and then integrate the product of that PDF and a to find $\mathbb{E}(D)$:

$$PDF = \frac{d}{da}CDF = \frac{d}{da}\left(\frac{a^2}{R^2}\right) = \frac{2a}{R^2}$$
$$\mathbb{E}(D) = \int_a a \cdot PDF da = \int_0^R a \cdot \left(\frac{2a}{R^2}\right) da = \left[\frac{2a^3}{3R^2}\right]_0^R = \frac{2}{3}\frac{R^3}{R^2} - 0 = \boxed{\frac{2R}{3}}.$$

$\square$

2. **A partial joint**. Define a joint density for X and Y as $f_{X,Y}(x,y) = \frac{12}{5} \cdot x \cdot (2-x-y)$ for $0 < x < 1$ and $0 < y < 1$, and zero outside of that region. Compute the conditional density of X given that $Y = y$.

Solution. Recall the conditional density formula:

$$f_{X,Y}(x,y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}.$$

First, let's find the marginal density of Y:

$$f_Y(y) = \int_x f_{X,Y}(x,y)\ dx = \int_0^1 \frac{12}{5}x(2-x-y)\ dx = \int_0^1 \left(\frac{24x}{5} - \frac{12x^2}{5} - \frac{12xy}{5}\right)\ dx$$
$$= \left[\frac{12x^2}{5} - \frac{4x^3}{5} - \frac{6x^2}{5}\right]_{x=0}^{x=1} = \frac{12-4-6y}{5} = \frac{8-6y}{5}.$$

Now, we can employ the formula for the conditional density:

$$
\begin{aligned}
f_{X|Y=y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} &= \frac{\left(\frac{12x}{5}\right)(2-x-y)}{\frac{8-6y}{5}} \\
&= \frac{24x - 12x^2 - 12xy}{8-6y} = \boxed{\frac{-6(x^2-2x+xy)}{4-3y}.}
\end{aligned}
$$

$\square$

3. **Stuck in the middle**. Suppose that a distribution is symmetric around a point c such that $f_X(c-x) = f_X(c+x)$ for all x. Show that the expected value X is c. Use this to argue that for a normal distribution $X \sim \mathcal{N}(\mu, \sigma^2)$ the expected value is μ.

SOLUTION. First, we want to show that $\mathbb{E}(X) = c$ for a symmetric distribution. Let's find a clever way to rewrite the definition of the expectation:

$$
\mathbb{E}(X) = \int_X X \cdot f_X(x)\ dx = \int_0^\infty \left((c-x)f_X(c-x) + (c+x)f_X(c+x)\right)\ dx.
$$

This effectively adds zero in a clever way, as we think of the integration process symmetrically. By offsetting the function by positive and negative x as done above, we capitalize on the fact that the distribution is symmetric about c. This then allows us to consider the integration process with only positive x values; instead of integrating across all x values, we can integrate from zero to infinity. We're essentially thinking of the integral as the sum of all slices x distance away from c on the x-axis, with their corresponding "heights" from the density function. For example, $(c-x)f_x(c-x) = \textit{distance} \cdot \textit{density} = \textit{point mass}$, so when we integrate we recover the expectation. Continuing with the calculus, we find

$$
\begin{aligned}
\mathbb{E}(X) &= \int_0^\infty \left((c-x)f_X(c-x) + (c+x)f_X(c+x)\right)\ dx \\
&= 2c\int_0^\infty f_X(c+x)\ dx \\
&= 2c \cdot \frac{1}{2} = c.
\end{aligned}
$$

Now, consider the probability density function of the normal distribution,

$$
f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}}
$$

Notice that the part of this equation that first depends on x can be rewritten as such:

$$
(x-\mu)^2 = (|x-\mu|)^2
$$

One can think of this expression as "the measure of the distance from μ squared." Thus, with respect to the density function, moving by x in the negative direction from μ produces the same function as moving by x in the positive direction from μ. Thus, the function is symmetrical about μ, and thus the expectation must be μ. $\square$

4. **Three short stories about MGFs.**

   (a) Show that the k^{th} derivative of $MGF_X(t)$ when $t = 0$ is the k^{th} moment of X.

   (b) Show that if $Y = a \cdot X$ then $MGF_Y(t) = MGF_X(at)$.

   (c) Show that if $X \perp Y$ then $MGF_{X+Y}(t) = MGF_X(t) \cdot MGF_Y(t)$.

SOLUTION. Consider the k^{th} derivative of the $MGF_X(t)$ when $t = 0$:

$$\frac{d^k}{dt^k} MGF_X(t) = \frac{d^k}{dt^k} \mathbb{E}(e^{tx}) = \frac{d^k}{dt^k} \int_x e^{tx} f_x(x)\ dx = \int_x \frac{d^k}{dt^k} e^{tx} f_x(x)\ dx.$$

To continue simplifying, let's expand e:

$$\begin{aligned}
\frac{d^k}{dt^k} MGF_X(t) &= \frac{d^k}{dt^k} e^{tk} \\
&= \frac{d^k}{dt^k}\left(1 + tx + \frac{t^2x^2}{2!} + \frac{t^3x^3}{3!} + \cdots + \frac{t^kx^k}{k!} + \frac{t^{k+1}x^{k+1}}{(k+1)!} + \cdots\right) \\
&= 0 + \cdots + x^k + tx^{k+1} + \cdots \quad \text{(thus when t=0, only } x^k \text{ remains).}
\end{aligned}$$

Continuing our original calculation, we find the k^{th} uncentered moment of X:

$$\frac{d^k}{dt^k} MGF_X(t) = \int_X x^k f_X(x)\ dx = \mathbb{E}(x^k).$$

As for when $Y = aX$, let's do a little LHS-RHS footwork to show $MGF_Y(t) = MGF_X(at)$:

$$LHS = MGF_Y(t) = \mathbb{E}(e^{ty}) = \mathbb{E}(e^{tax}) = MGF_X(at) = RHS.$$

Finally, when $X \perp Y$ then $MGF_{X+Y}(t) = MGF_X(t) \cdot MGF_Y(t)$:

$$LHS = MGF_{X+Y}(t) = \mathbb{E}(e^{t(x+y)}) = \mathbb{E}(e^{tx} \cdot e^{ty}).$$

We also know from class that when $X \perp Y$, $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$. So,

$$LHS = \mathbb{E}(e^{tx} \cdot e^{ty}) = \mathbb{E}(e^{tx})\mathbb{E}(e^{ty}) = MGF_X(t)MGF_Y(t) = RHS. \qquad \square$$

5. **MGFs, for example.**

   (a) Find the MGF of a geometric random variable and use it to find the mean and the variance

   (b) The MGF of a random variable uniquely identifies it. Use the properties we know to show that if $X \sim \text{Poisson}(\lambda)$ and $Y \sim \text{Poisson}(\mu)$ are independent, then $X+Y \sim \text{Poisson}(\mu+\lambda)$.

SOLUTION. We know the geometric random variable to be discrete, so let's employ the discrete formula for expectation:

$$\begin{aligned}
MGF_X(t) = \mathbb{E}(e^{tx}) &= \sum_k e^{tk}(1-p)^{k-1}p \\
&= \sum_k e^{tk}(1-p)^{k-1}p \cdot \left(\frac{1-p}{1-p}\right) = \frac{p}{1-p}\sum_k (e^t(1-p))^k.
\end{aligned}$$

Note that this follows the form of a geometric series with $r = e^t(1-p)$. So, applying the geometric formula for a finite sum,

$$\sum_{k=0}^{\infty} r^k = \frac{1}{1-r},$$

we find that

$$MGF_X = \frac{p}{1-p}\left(\frac{e^t(1-p)}{1-e^t(1-p)}\right) = \boxed{\frac{pe^t}{1-e^t(1-p)}.}$$

Now, we take the first derivative with respect to t and let $t = 0$ to find the first uncentered moment, or $\mathbb{E}(X)$. Applying the quotient rule, we find that

$$\begin{aligned}\frac{d}{dt}MGF_X(t) &= \frac{(1-e^t(1-p))(pe^t) - (pe^t)(-e^t(1-p))}{(1-e^t(1-p))^2} \\ &= \frac{(pe^t)(-e^t(1-p) + e^t(1-p)}{(1-e^t(1-p))^2},\end{aligned}$$

which when $t = 0$, gives $\boxed{\mu = \frac{1}{p},}$ which we know to be the expectation of a geometric random variable with weight p.

Next, differentiate again to find the second uncentered moment. After yet another charming quotient rule, we find

$$\mathbb{E}(X^2) = \frac{(p^2)(p) - (p)(2p(p-1))}{(p^2)^2} = \frac{-p+2}{p^2}.$$

Finally, employing the equation for variance, we see

$$\begin{aligned}\mathbb{V}(X) &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 \\ &= \frac{-p+2}{p^2} - \left(\frac{1}{p}\right)^2 = \boxed{\frac{1-p}{p^2} = \sigma^2.}\end{aligned}$$

Now consider $X \sim \text{Poisson}(\lambda) \perp Y \sim \text{Poisson}(\mu)$ and $X + Y \sim \text{Poisson}(\mu + \lambda)$. Recall from Problem 4 that if $X \perp Y$, then $MGF_{X+Y}(t) = MGF_X(t) \cdot MGF_Y(t)$. So,

$$\begin{aligned}MGF_{X+Y}(t) &= MGF_X(t) \cdot MGF_Y(t) \\ &= e^{\lambda(e^t-1)} \cdot e^{\mu(e^t-1)} \\ &= e^{(\mu+\lambda)(e^t-1)}, \quad \text{which shows that } X + Y \sim \text{Poisson}(\mu + \lambda). \qquad \square\end{aligned}$$

6. **Independent Covariation**. Find a counter-example to the claim that zero covariance implies independence.

SOLUTION. Let X be the random variable that measures the amount of peanuts I give Connor during the experiment wherein I give Connor six peanuts. Let Y be the random variable that measures the amount of peanuts The Feds give Connor during the experiment wherein The Feds give Connor $2X$ peanuts. The quantity of peanuts that Connor receives from The Feds is completely dependent on the quantity of peanuts that I give him, but it's trivial to see that X and Y each have zero variance, and that the covariance between them must also be zero. $\square$

7. **Covariation in mean**. Let $X_1, \ldots, X_N$ be independent and identically distributed random variables with mean μ and variance σ^2. Show that $\mathrm{Cov}(X_i - \overline{X}, X) = 0$.

Solution. Recall that,

$$\overline{X} = \frac{1}{N}\sum_{i=1}^{N} X_i, \quad \mathrm{Cov}(XY) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y), \quad \text{and} \quad \mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2.$$

Basic manipulation of the covariance equation with the relevant random variables yields

$$\mathrm{Cov}(X_i - \overline{X}, X) = \mathbb{E}(\overline{X}X_i) - \mathbb{E}(\overline{X})^2 + \mathbb{E}(X_i - \overline{X})\mathbb{E}(\overline{X}),$$

which equals $\mathrm{Cov}(X_i - \overline{X}, X) = \mathbb{E}(\overline{X}X_i) - \mathbb{E}(\overline{X})^2$, because the third term cancels to zero. Let's consider the term $\mathbb{E}(\overline{X}X_i)$:

$$\begin{aligned} \mathbb{E}(\overline{X}X_i) &= \mathbb{E}\left(X_i \frac{1}{N}\sum_{j=1}^{N} X_j\right) = X_i \frac{1}{N}\sum_{j=1}^{N}\mathbb{E}(X_j X_i) \\ &= \frac{(N-1)\mathbb{E}(X_j)\mathbb{E}(X_i) + (1)\mathbb{E}(X_j)\mathbb{E}(X_i)}{N}. \end{aligned}$$

(Note, there are $(N-1)$ terms where $i \neq j$ and 1 term where $i = j$.)

$$\begin{aligned} \mathbb{E}(\overline{X}X_i) &= \frac{1}{N}\left[(N-1)\mu^2 + \mathbb{E}(X_i)^2\right] \\ &= \frac{1}{N}\left[(N-1)\mu^2 + \mathbb{V}(X_i) + \mathbb{E}(X_i^2)\right] \quad \text{(variance equation)} \\ &= \mu^2 + \frac{\sigma^2}{N}. \end{aligned}$$

Next, let's analyze $\mathbb{E}(\overline{X}^2)$. Again, from the variance formula we find $\mathbb{E}(\overline{X}^2) = \mathbb{V}(\overline{X}) + \mathbb{E}(\overline{X})^2$, and using the result derived in class, namely that $\mathbb{V}(\overline{X}_N) = \frac{\sigma^2}{N}$, we can see that

$$\mathbb{E}(\overline{X}^2) = \frac{\sigma^2}{N} + \mu^2.$$

Combining our newly analyzed terms, we find that

$$\begin{aligned} \mathrm{Cov}(X_i - \overline{X}, X) &= \mathbb{E}(\overline{X}X_i) - \mathbb{E}(\overline{X})^2 \\ &= \left(\frac{\sigma^2}{N} + \mu^2\right) - \left(\frac{\sigma^2}{N} + \mu^2\right) \\ &= 0. \end{aligned}$$

$\square$

8. **Indicators of covariation.** Suppose $\mathbf{1}_A$and $\mathbf{1}_B$ are indicator random variables for events A and B. Find Cov($\mathbf{1}_A$,$\mathbf{1}_B$).

SOLUTION. Applying the covariance formula, we see that

$$\mathrm{Cov}(\mathbf{1}_A, \mathbf{1}_B) = \mathbb{E}(\mathbf{1}_A \mathbf{1}_B) - \mathbb{E}(\mathbf{1}_A)\mathbb{E}(\mathbf{1}_B).$$

Recall, $\mathbb{E}(\mathbf{1}_A) = \sum_{k=0}^{1} k \cdot \mathbb{P}(\mathbf{1}_A = k) = \mathbb{P}(\mathbf{1}_A = 1) = \mathbb{P}(A)$. Also note that the only time the product of two indicators is 1 (as opposed to zero) is when both events are true. Using these facts we can see that

$$\begin{aligned}
\mathrm{Cov}(\mathbf{1}_A, \mathbf{1}_B) &= \mathbb{E}(\mathbf{1}_{A\cap B}) - \mathbb{P}(\mathbf{1}_A)\mathbb{P}(\mathbf{1}_B) \\
&= \mathbb{P}(A \cap B) - \mathbb{P}(A)]P(B) \\
\mathrm{Cov}(\mathbf{1}_A, \mathbf{1}_B) &= \boxed{\mathbb{P}(A|B)\mathbb{P}(B) - \mathbb{P}(A)P(B).}
\end{aligned}$$

□

9. **Covarying opinion.** Suppose a population of N individuals of whom $n \leq N$ are sampled. Define $\mathbf{1}_i$ to be an indicator on the event that i is included in the same or not. Find the the expectation and variance of $\mathbf{1}_i$, along with $\mathrm{Cov}(\mathbf{1}_i, \mathbf{1}_j)$.

SOLUTION. We immediately see that $\mathbb{E}(\mathbf{1}_i) = \mathbb{P}(i) = \boxed{\frac{n}{N},}$ as a classic Cardano probability. As for the variance, we turn to the formula,

$$\mathbb{V}(\mathbf{1}_i) = \mathbb{E}(\mathbf{1}_i^2) - \mathbb{E}(\mathbf{1}_i)^2.$$

Notice that $\mathbb{E}(\mathbf{1}_i^2) = \sum_{k=0}^{1} k^2 \cdot \mathbb{P}(\mathbf{1}_i = 1) = \mathbb{P}(\mathbf{1}_i = 1) = \frac{n}{N}$. Hence,

$$\mathbb{V}(\mathbf{1}_i) = \frac{n}{N} - \left(\frac{n}{N}\right)^2 = \boxed{\frac{n(N-n)}{N^2}.}$$

Finally, let's inspect the covariance equation:

$$\begin{aligned}
\mathrm{Cov}(\mathbf{1}_i, \mathbf{1}_j) &= \mathbb{E}(\mathbf{1}_i \mathbf{1}_j) - \mathbb{E}(\mathbf{1}_i)\mathbb{E}(\mathbf{1}_j) \\
&= \mathbb{P}(\mathbf{1}_{i\cap j}) - \frac{n^2}{N^2} \\
&= \frac{\binom{2}{2}\binom{N-2}{n-2}}{\binom{N}{n}} - \frac{n^2}{N^2} \\
&= \boxed{\frac{n^2-n}{N^2-N} - \frac{n^2}{N^2}.}
\end{aligned}$$

□