

Problem Set #2

Danny Edgel
Econ 709: Economic Statistics and Econometrics I
Fall 2020

September 20, 2020

Collaborated with Sarah Bass, Emily Case, Michael Nattinger, and Alex Von Hafften

Question 1

Suppose that $Y = X^3$ and $f_X(x) = 42x^5(1-x)$, $x \in (0, 1)$. Find the PDF of Y , and show that the PDF integrates to 1.

We know that the CDF of Y , $F_Y(y)$ is equal to $F_X(f^{-1}(x))$. So we can solve for the PDF of Y by first finding its CDF:

$$\begin{aligned}f^{-1}(y) &= \sqrt[3]{y} \\F_X(x) &= \int_0^x 42t^5(1-t)dt = 42 \int_0^x t^5 - t^6 dt = 42\left(\frac{1}{6}x^6 - \frac{1}{7}x^7\right) \\F_X(f^{-1}(y)) &= 42\left(\frac{1}{6}y^{6/3} - \frac{1}{7}y^{7/3}\right) = 42y^2\left(\frac{1}{6} - \sqrt[3]{y}\right) = F_Y(y) \\f_Y(y) &= \frac{d}{dy}F_Y(y) = 14y - 14y\sqrt[3]{y}\end{aligned}$$

Since we already know $F_Y(y)$, we can easily show that the PDF of Y integrates to 1:

$$\int_0^1 f_Y(y)dy = F_Y(1) - F_Y(0) = 1^2(7 - 6\sqrt[3]{1}) - 0 = 7 - 6 = 1$$

Question 2

Consider the CDF $F_X(x) = \begin{cases} 1.2x & \text{if } x \in [0, 0.5) \\ 0.2 + 0.8x & \text{if } x \in [0.5, 1] \end{cases}$, and the function

$$f_X(x) = \begin{cases} 1.2 & \text{if } x \in [0, 0.5) \\ a & \text{if } x = 0.5 \\ 0.8 & \text{if } x \in (0.5, 1] \end{cases}$$

Show that f_X is the density function of F_X as long as $a \geq 0$. That is, show that for all $x \in [0, 1]$, $F_X(x) = \int_0^x f_X(t)dt$.

We can define $F_X(x) = \int_0^x f_X(t)dt$ on a case-by-case basis:

$$\begin{aligned} x \in [0, 0.5) : \int_0^x f_X(t)dt &= \int_0^x 1.2dt = [1.2t]_0^x = 1.2x \\ x = 0.5 : \int_0^x f_X(t)dt &= \int_0^0 1.2dt + \int_0^x a dt = [1.2t]_0^0 + [at]_0^x = 0 + ax = 0.5a \\ &= 1.2(0.5) - 0 + ax - 0.5a = 0.6 + 0.5x - 0.5x = 0.6 \\ x \in (0.5, 1) : \int_0^x f_X(t)dt &= \int_0^0 1.2dt + \int_0^x a dt + \int_0^x 0.8dt = [1.2t]_0^0 + [at]_0^x + [0.8t]_0^x \\ &= 0.6 + 0.8x - 0.4 = 0.8x + 0.2 \end{aligned}$$

Since $0.6 = 1.2x$ when $x = 0.5$, then $\int_0^x 1.2dt = \int_0^0 1.2dt + \int_0^x a dt$ when $x = 0.5$. Thus, $\forall x \in [0, 1]$, $F_X(x) = \int_0^x f_X(t)dt$.

Question 3

Let X have the PDF $f_X(x) = \frac{2}{9}(x+1)$, $x \in [-1, 2]$. Find the PDF of $Y = X^2$. Note that this is a bit different from the exercise in the lecture note.

Given $f_X(x)$ and $Y = g(X) = X^2$, we can solve:

$$f_Y(y) = \begin{cases} f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right|, & y \in Y \\ 0, & y \notin Y \end{cases}$$

Solving for each element of the first case of $f_Y(y)$ yields:

$$\begin{aligned} f_X(g^{-1}(y)) &= \frac{2}{9} \\ \frac{d}{dy} g^{-1}(y) &= 1/(2\sqrt{y}) = \frac{1}{2\sqrt{y}} \end{aligned}$$

Since $x \in [-1, 2]$ and $\frac{d}{dy}g^{-1}(y) = 1/(2\sqrt{y})$, our first case of $f_Y(y)$ is defined on $y \in (0, 4]$. Thus,

$$f_Y(y) = \begin{cases} \frac{1}{18\sqrt{y}}, & y \in (0, 4] \\ 0, & y \leq 0 \vee y > 4 \end{cases}$$

Question 4

A median of a distribution is a value m such that $P(X \leq m) \geq \frac{1}{2}$ and $P(X \geq m) \geq \frac{1}{2}$. Find the median of the distribution $f(x) = 1/\pi(1+x^2)$, $x \in \mathbb{R}$.

Given that $P(X \leq x) = F_X(x)$, we need to find m such that $\frac{1}{2} \leq F_X(m) \leq \frac{1}{2}$:

$$\begin{aligned} \frac{1}{2} &\leq \frac{1}{\sqrt{\pi}} \int_{-\infty}^m f_X(x) \leq \frac{1}{2} \\ \frac{\sqrt{\pi}}{2} &\leq \int_{-\infty}^m \frac{1}{\pi}(1+t^2)dt \leq \frac{\sqrt{\pi}}{2} \\ \frac{\sqrt{\pi}}{2} &\leq [\arctan(x/\sqrt{\pi})]_{-\infty}^m \leq \frac{\sqrt{\pi}}{2} \\ \frac{\sqrt{\pi}}{2} &\leq (\arctan(m/\sqrt{\pi}) - 1) \leq \frac{\sqrt{\pi}}{2} \\ \tan(\sqrt{\pi}/2) &\leq (m/\sqrt{\pi}) - 1 \leq \tan(\sqrt{\pi}/2) \end{aligned}$$

$$\therefore m = \sqrt{\pi}(1 + \tan(\sqrt{\pi}/2)) \blacksquare$$

Question 5

Show that if X is a continuous random variable, then $\min_a E|X - a| = E|X - m|$, where m is the median of X .

We simply need to find a that minimizes $E(X - a)$. Then:

$$\begin{aligned} E(X - a) &= \int_{-\infty}^{\infty} f_X(t)|t - a|dt = \int_{-\infty}^a f_X(t)(a - t)dt + \int_a^{\infty} f_X(t)(t - a)dt \\ \frac{d}{da}E(X - a) &= \int_{-\infty}^a f_X(t)dt - \int_a^{\infty} f_X(t)dt = F_X(a) - (1 - F_X(a)) = 0 \\ F_X(a) &= \frac{1}{2} \end{aligned}$$

$$\therefore a = m \blacksquare$$

Question 6

Let μ_n denote the n th central moment of a random variable X . Two quantities of interest, in addition to the mean and variance are

$$\alpha_3 = \frac{\mu_3}{\mu_2^{3/2}} \text{ and } \alpha_4 = \frac{\mu_4}{\mu_2^2}$$

The value α_3 is called the skewness and α_4 is called the kurtosis. The skewness measures the lack of symmetry in the density function. The kurtosis measures the peakedness or flatness of the density function.

1. Show that if a density function is symmetric about a point a , then $\alpha_3 = 0$

Let X be a continuous random variable that is symmetrically distributed about some point a . Thus, $E(X) = a$. Now, define $Y = X - a$ and $g(y) = y^3$. Then, we can derive:

$$\begin{aligned} \mu_3 &= E(Y^3) \\ &= \int_{-\infty}^a f_X(y)g(y)dy + \int_a^{\infty} f_X(y)g(y)dy \\ &= \int_a^{\infty} f_X(-y)g(-y)dy + \int_a^{\infty} f_X(y)g(y)dy \\ &= -\int_a^{\infty} f_X(y)g(y)dy + \int_a^{\infty} f_X(y)g(y)dy \\ \mu_3 &= 0 \end{aligned}$$

Therefore, $\alpha_3 = \frac{\mu_3}{\mu_2^{3/2}} = 0$.

2. Calculate α_3 for $f(x) = e^{-x}$, $x \geq 0$, a density function that is skewed to the right.

First, we must solve for the mean and second and third central moments:

$$\begin{aligned} E(X) &= \int_0^{\infty} e^{-t}t dt = [-te^{-t}]_0^{\infty} - \int_0^{\infty} e^{-t} dt = -0 + 0 - [e^{-x}]_0^{\infty} = 1 \\ \mu_2 &= E(X^2) - 1^2 = \int_0^{\infty} e^{-t}t^2 dt - 1 = -2 \int_0^{\infty} e^{-t}t dt - 1 = 2 - 1 = 1 \\ \mu_3 &= \int_0^{\infty} e^{-t}(t-1)^3 dt = \int_0^{\infty} e^{-t}(t^3 - 3t^2 + 3t - 1) dt = E(X^3) - 3E(X^2) + 3E(X) - 1 \\ E(X^3) &= \int_0^{\infty} e^{-t}t^3 dt = 0 - 3 \int_0^{\infty} t^2 e^{-t} dt = 3(2) = 6 \\ \mu_3 &= 6 - 3(2) + 3 - 1 = 2 \end{aligned}$$

Thus,

$$\alpha_3 = \frac{\mu_3}{\mu_2^{3/2}} = \frac{2}{1} = 2$$

3. Calculate α_4 for the following density functions and comment on the peakedness of each:

(a) $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, x \in \mathbb{R}$

f is simply the normal distribution, so we know that it has moment-generating function $M(t) = e^{t^2/2}$. Then, we can calculate the following derivatives of M :

$$\begin{aligned} M'(t) &= te^{t^2/2} \\ M^{(2)}(t) &= (t^2 + 1)e^{t^2/2} \\ M^{(3)}(t) &= (3t + t^3)e^{t^2/2} \\ M^{(4)}(t) &= (3t^3 + 6t + 3)e^{t^2/2} \end{aligned}$$

Using $\mu_k = M^{(k)}(0)$, we can solve for α_4 :

$$\alpha_4 = \frac{\mu_4}{\mu_2^2} = \frac{M^{(4)}(0)}{(M^{(2)}(0))^2} = \frac{3}{(-1)^2} = 3$$

A kurtosis of 3 is very low, indicating that the tails of the distribution are fairly skinny, with most of the data falling near the mean, where there is a "peak".

(b) $f(x) = 1/2, x \in (-1, 1)$

We can begin by finding $E(X)$, μ_2 , and μ_4 . This distribution is symmetric about 0, so $E(X) = 0$. Then,

$$\begin{aligned} \mu_2 &= \int_{-1}^1 \frac{1}{2} t^2 dt = \left[\frac{1}{6} t^3 \right]_{-1}^1 = \frac{1}{6} (1 + 1) = \frac{1}{3} \\ \mu_4 &= \int_{-1}^1 \frac{1}{2} t^4 dt = \left[\frac{1}{10} t^5 \right]_{-1}^1 = \frac{1}{10} (1 + 1) = \frac{1}{5} \end{aligned}$$

Thus,

$$\alpha_4 = \frac{\mu_4}{\mu_2^2} = \frac{1/5}{(1/3)^2} = \frac{9}{5}$$

The kurtosis of this function is even smaller than that of the normal distribution, indicating a high peakedness of the distribution, even though this distribution is flat. This is because the values only fall from -1 to 1 , whereas most others fall in an infinite range. The density from -1 to 1 , then, can be seen as a sharp, steep peak with a flat top.

(c) $f(x) = \frac{1}{2} e^{-|x|}, x \in \mathbb{R}$

This distribution is clearly symmetric about 0, so $E(X) = 0$. Then,

$$\begin{aligned}
\mu_2 &= \int_{-\infty}^0 \frac{1}{2} e^t t^2 dt + \int_0^{\infty} \frac{1}{2} e^{-t} t^2 dt \\
&= \left[\frac{1}{2} t^2 e^t \right]_{-\infty}^0 + \int_{-\infty}^0 t e^t + \left[\frac{1}{2} t^2 e^{-t} \right]_0^{\infty} + \int_0^{\infty} t e^{-t} \\
&= [t e^t]_{-\infty}^0 + \int_{-\infty}^0 e^t + [t e^{-t}]_0^{\infty} + \int_0^{\infty} e^{-t} \\
&= [e^t]_{-\infty}^0 - [e^{-t}]_0^{\infty} = 1 - 0 - 0 + 1 \\
\mu_2 &= 2
\end{aligned}$$

And, simplifying intermediate steps where the definite integral is known to equal zero by the above steps:

$$\begin{aligned}
\mu_4 &= \int_{-\infty}^0 \frac{1}{2} e^t t^4 dt + \int_0^{\infty} \frac{1}{2} e^{-t} t^4 dt \\
&= 0 + \int_{-\infty}^0 2 e^t t^3 dt - 0 + \int_0^{\infty} 2 e^{-t} t^3 dt \\
&= 0 + \int_{-\infty}^0 6 e^t t^2 dt - 0 + \int_0^{\infty} 6 e^{-t} t^2 dt \\
&= 0 + \int_{-\infty}^0 12 e^t t dt - 0 + \int_0^{\infty} 12 e^{-t} t dt \\
&= 0 + \int_{-\infty}^0 12 e^t dt - 0 + \int_0^{\infty} 12 e^{-t} dt \\
&= [12 e^t]_{-\infty}^0 - [12 e^{-t}]_0^{\infty} = 12 - 0 - 0 + 12 \\
\mu_4 &= 24
\end{aligned}$$

Thus,

$$\alpha_4 = \frac{\mu_4}{\mu_2^2} = \frac{24}{2^2} = 6$$

Therefore, this distribution is half as peaked as the normal distribution.