(Due Tuesday 04/09/2019 right before the class)
&
(Your homework shall be stapled if it contains multiple pages.)

SPRING/2019/MA526: HOMEWORK 8

Instructor: Guangqu Zheng[1]; Grader: Chessa Mccalla[2]

Total points: 20

Q1 (4 pt) Let X be a standard normal random variable, find the expected value of $|X|$, i.e. compute

$$\mathbb{E}\big[|X|\big] .$$

And find the probability density function of the random variable $|X|$.

Ans: By direct computation,

$$\mathbb{E}\big[|X|\big] = \int_{-\infty}^{\infty} |x| \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx = 2\int_{0}^{\infty} |x| \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx$$

here we used the symmetry in the last equality. Now we compute the integral

$$\int_0^\infty |x| e^{-x^2/2}\, dx = \int_0^\infty x e^{-x^2/2}\, dx = -\int_0^\infty de^{-x^2/2} = -e^{-x^2/2}\big|_{x=0}^{x=\infty} = 1$$

so that

$$\mathbb{E}\big[|X|\big] = 2\frac{1}{\sqrt{2\pi}} \int_0^\infty |x| e^{-x^2/2}\, dx = \sqrt{2/\pi}$$

Now we look for the PDF of $Y := |X|$. Clearly, Y is a nonnegative random variable (Remember this!). For $y > 0$, we have

$$F_Y(y) = \mathbb{P}\big(Y \le y\big) = \mathbb{P}\big(|X| \le y\big) = \mathbb{P}\big(-y \le X \le y\big) = \Phi(y) - \Phi(-y) .$$

Here we use Φ and ϕ to denote the CDF, PDF of standard normal respectively. So that the PDF of Y can be found by looking at the derivative of $F_Y(y)$:

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy}\Big(\Phi(y) - \Phi(-y)\Big) = \phi(y) + \phi(-y)$$

since $\frac{d}{dy}\Phi(y) = \phi(y)$ and $\frac{d}{dy}\Phi(-y) = -\phi(-y)$. We know the expression for $\phi(y)$, then we write

$$f_Y(y) = \frac{2}{\sqrt{2\pi}} e^{-y^2/2} = \sqrt{2/\pi} e^{-y^2/2} , \quad for \quad y > 0$$

meanwhile $f_Y(y) = 0$ for $y \le 0$. This gives us the PDF of Y:

$$f_Y(y) = \begin{cases} \sqrt{2/\pi}\; e^{-y^2/2} , & for \quad y > 0 \\ 0 , & for \quad y \le 0 \end{cases}$$

[1]gzheng90@ku.edu; Office hours: TuTh 11:00-11:50; Office = 641 Snow Hall
[2] chessa_m@ku.edu

Q2 (4pt) Gauges are used to reject all components for which a certain dimension is not within the specification

$$1.50 \pm d.$$

It is known that this measurement is normally distributed with mean 1.5 and standard deviation 0.2. Determine the value of d such that these specifications "cover" 99.8%. [Use the normal table.]

Explain why the assumption that "**this measurement is normally distributed** " is not ridiculous at all? (We know that the normal random variable could take negative values.)

Ans: We can view the measurement as the normal random variable given by $M := 1.5 + 0.2Z$ with Z standard normal. The specification $1.50 \pm d$ is equivalent to $-d \leq 0.2Z \leq d$ or

$$-5d \leq Z \leq 5d\,.$$

The event that these specifications "cover" 99.8% can be expressed as $\mathbb{P}\big(1.5 - d \leq M \leq 1.5 + d\big) = 99.8\%$. By previous discussion, we are looking for d such that

$$\mathbb{P}\big(-5d \leq Z \leq 5d\big) = 99.8\%$$

thus, we are looking for d such that

$$\mathbb{P}\big(Z \leq 5d\big) = 99.9\%.$$

In fact, there are 0.2% area left outside the region $[-5d, 5d]$, and by symmetry of the standard normal density function, we know that there is 0.1% area to the right of $5d$. This tells us $\mathbb{P}(Z \leq 5d) = 99.9\%$

From the normal table, we find that $5d = 3.1$, so $d = 0.62$.

The assumption that "**this measurement is normally distributed** " is not ridiculous at all, because as we see from the previous result, there is very small chance for M to be outside 1.5 ± 0.62. moreover, as one can see also from the normal table, $\mathbb{P}\big(M \leq 0\big) \approx 0$.

Q3 (4pt) One-sixth of the male freshmen entering a large state school are out-of-state students. If the students are assigned at random to dormitories, 180 to a building, what is the probability that in a given dormitory at least one-fifth of the students are out-of-state? [Use normal approximation.]

Ans: Let X be a Binomial$(180, 1/6)$ random variable, then its mean is 30 and variance is 25 , the standard deviation 5. The desired probability is

$$\mathbb{P}\big(X \geq \frac{180}{5}\big) = \mathbb{P}\big(X \geq 36\big) = \mathbb{P}\big(\frac{X-30}{5} \geq \frac{36-30}{5}\big) = \mathbb{P}\big(\frac{X-30}{5} \geq 1.2\big) \approx \mathbb{P}\big(Z \geq 1.2\big)$$

with Z standard normal. From the normal table, $\mathbb{P}\big(Z \leq 1.2\big) = 0.8849$ so the desired probability is close to 0.1151.

you can also consider the **continuity correction** as follows:

$$\mathbb{P}\big(X \geq \frac{180}{5}\big) = 1 - \mathbb{P}\big(X \leq 35\big)$$

and

$$\mathbb{P}\big(X \leq 35\big) \approx \mathbb{P}\big(Z \leq \frac{35 + 0.5 - 30}{5}\big) = \mathbb{P}\big(Z \leq 1.1\big) \approx 0.8643$$

so $\mathbb{P}\big(X \geq \frac{180}{5}\big) = 1 - \mathbb{P}\big(X \leq 35\big) \approx 0.1357$

Q4 (2+2pt) Consider X a standard exponential random variable, that is, X has the following probability density function

$$f(x) = e^{-x} \quad x \geq 0 \quad and \quad f(x) = 0 \quad x < 0.$$

(1) Find the cumulative distribution function of X. Let us denote it by F.

(2) What is the distribution of $F(X)$? Explain with enough details to get your points here.

Ans:

$$F(x) = 1 - e^{-x} \quad x \geq 0 \quad and \quad F(x) = 0 \quad x < 0.$$

Now $F(X) = 1 - e^{-X}$. It is clear that $0 \leq F(X) \leq 1$, because F is a CDF. $F(X)$ is a random variable with values in $[0, 1]$. Suppose $0 < t < 1$, we have

$$\mathbb{P}\big(F(X) \leq t\big) = \mathbb{P}\big(1 - e^{-X} \leq t\big) = \mathbb{P}\big(1 - t \leq e^{-X}\big) = \mathbb{P}\big(\log_e(1-t) \leq -X\big) = \mathbb{P}\big(-\log_e(1-t) \geq X\big)$$

which is equal to $F\big(-\log_e(1-t)\big)$.

For $0 < t < 1$,

$$F\Big(\log_e \frac{1}{1-t}\Big) = 1 - e^{-y} \quad with \quad y = -\log_e(1-t)$$

$e^{-y} = 1 - t$, so $F\big(\log_e \frac{1}{1-t}\big) = 1 - (1-t) = t$. Hence, if $0 < t < 1$,

$$\mathbb{P}\big(F(X) \leq t\big) = t$$

this is exactly the CDF of uniform random variable on $(0, 1)$. To conclude, $F(X) \sim \text{Uniform}(0, 1)$.

Q5 (4pt) A certain type of device has an advertised failure rate of 0.01 per hour. The failure rate is constant and the exponential distribution applies.

(1) What is the mean time to failure?

(2) What is the probability that 200 hours will pass before a failure is observed?

Explain your answers with enough details to get your points here.

Ans: mean time to failure is 100 hours, and we consider the exponential random variable T with PDF given by

$$f(x) = 0.01e^{-0.01x} \quad for \quad x > 0 \quad and \quad f(x) = 0 \quad for \quad x \leq 0\,.$$

And the probability that 200 hours will pass before a failure is observed is

$$\mathbb{P}(T > 200) = \int_{200}^{\infty} 0.01e^{-0.01x}dx = 1 - F(200)$$

F is the CDF of this exponential random variable, and we know its expression from lecture or you can compute it:

$$F(x) = 1 - e^{-0.01x} \quad for \quad x > 0 \quad and \quad F(x) = 0 \quad for \quad x \leq 0\,.$$

so $F(200) = 1 - e^{-2}$ and therefore

$$\mathbb{P}(T > 200) = 1 - F(200) = e^{-2} \approx 0.13533528323\,.$$