Tutorial 4

October 15, 2020

Question 1

Let U_1 and U_2 be two independent Uniform$[0,1]$ random variables, and let $X = \min(U_1, U_2)$ be the minimum between them. Show that the density of X is

$$f_X(x) = 2 - 2x, \quad 0 \le x \le 1$$

[Hint: Start by computing $F_X(x) = \mathbf{P}(X \le x)$.]

$$\begin{aligned}
\mathbf{P}(X \le x) &= 1 - \mathbf{P}(X > x) \\
&= 1 - \mathbf{P}(U_1 > x \cap U_2 > x) && \text{(Definition of minimum)} \\
&= 1 - (\mathbf{P}(U_1 > x) \cdot \mathbf{P}(U_2 > x)) && \text{(Independence)} \\
&= 1 - (\mathbf{P}(U > x))^2 && \text{(Identically distributed)} \\
&= 1 - (1 - \mathbf{P}(U \le x))^2 \\
&= 1 - (1 - x)^2 \\
&= 1 - (1 - 2x + x^2) \\
&= 2x - x^2
\end{aligned}$$

$$F_X(x) = \begin{cases} 0 & x < 0 \\ 2x - x^2 & 0 \le x \le 1 \\ 1 & x > 1 \end{cases}$$

$$\begin{aligned}
f_X(x) &= \frac{d}{dx} F_X(x) \\
&= \begin{cases} 2 - 2x & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}
\end{aligned}$$

Question 2

A die is rolled 24 times. Let S_{24} represent the sum of the 24 numbers rolled. Using the central limit theorem, approximate $\mathbf{P}(S_{24} \geq 100)$.

Let X be the outcome of a single roll of a die. Then its expected value is:

$$\mathbf{E}(X) = \sum_{x=1}^{6} x \cdot p(x) = \frac{1}{6}\sum_{x=1}^{6} x = \frac{1}{6} \cdot \frac{6(6+1)}{2} = \frac{7}{2} = 3.5$$

Similarly,

$$\mathbf{E}\left(X^2\right) = \sum_{x=1}^{6} x^2 \cdot p(x) = \frac{1}{6}\sum_{x=1}^{6} x^2 = \frac{1}{6} \cdot \frac{6(6+1)(2(6)+1)}{6} = \frac{91}{6}$$

The variance is computed as:

$$\mathbf{Var}(X) = \mathbf{E}\left(X^2\right) - (\mathbf{E}(X))^2 = \frac{91}{6} - (3.5)^2 = \frac{35}{12}$$

Let S_{24} represent the sum of the 24 rolled numbers, i.e.

$$S_{24} = X_1 + X_2 + \ldots + X_{24}$$

Then its expected value is:

$$\begin{aligned}
\mathbf{E}(S_{24}) &= \mathbf{E}(X_1 + X_2 + \ldots + X_{24}) \\
&= \mathbf{E}(X_1) + \mathbf{E}(X_2) + \ldots + \mathbf{E}(X_{24}) \\
&= \sum_{i=1}^{24} \mathbf{E}(X_i) \\
&= 24 \cdot \mathbf{E}(X) \\
&= 24 \cdot \frac{7}{2} \\
&= 84
\end{aligned}$$

Assuming each roll is independent of one another, its variance is:

$$\begin{aligned}
\mathbf{Var}(S_{24}) &= \mathbf{Var}(X_1 + X_2 + \ldots + X_{24}) \\
&= \mathbf{Var}(X_1) + \mathbf{Var}(X_2) + \ldots + \mathbf{Var}(X_{24}) \\
&= \sum_{i=1}^{24} \mathbf{Var}(X_i) \\
&= 24 \cdot \mathbf{Var}(X) \\
&= 24 \cdot \frac{35}{12} \\
&= 70
\end{aligned}$$

Applying the central limit theorem, let

$$Z := \frac{S_{24} - 84}{\sqrt{70}} \;\dot{\sim}\; N(0,1)$$

Since we are going from a discrete distribution to a continuous distribution, we need to make a continuity correction. We also note that for a continuous distribution, $\mathbf{P}(X < x) = \mathbf{P}(X \le x)$.

$$\begin{aligned}
\mathbf{P}(S_k \ge 100) &\approx \mathbf{P}(S_k \ge 100 - 0.5) \\
&= \mathbf{P}\left(\frac{S_{24} - 84}{\sqrt{70}} \ge \frac{99.5 - 84}{\sqrt{70}}\right) \\
&= \mathbf{P}\left(Z \ge \frac{15.5}{\sqrt{70}}\right) \\
&= 1 - \mathbf{P}\left(Z \le \frac{15.5}{\sqrt{70}}\right) \\
&= 0.03196954
\end{aligned}$$

```
R> pnorm(15.5/sqrt(70), lower.tail=FALSE)
```

Compared to the exact answer, 0.031760, this approximation is quite close. Getting the exact answer requires a bit of work though.

Question 3

A die is rolled k times. Let S_k represent the sum of the k numbers rolled. Using the central limit theorem, how large should k be so that $\mathbf{P}(S_k \ge 100) > 0.05$?

Let

$$S_k = X_1 + X_2 + \ldots + X_k$$

Then

$$\mathbf{E}(S_k) = k \cdot \frac{7}{2}, \quad \mathbf{Var}(S_k) = k \cdot \frac{35}{12}$$

Applying the central limit theorem, let

$$Z := \frac{S_k - (7k/2)}{\sqrt{35k/12}} \;\dot{\sim}\; N(0,1)$$

$$\begin{aligned}
\mathbf{P}(S_k \ge 100) &\approx \mathbf{P}(S_k \ge 99.5) \\
&= \mathbf{P}\left(\frac{S_k - (7k/2)}{\sqrt{35k/12}} \ge \frac{99.5 - (7k/2)}{\sqrt{35k/12}}\right) \\
&= \mathbf{P}\left(Z \ge \frac{99.5 - (7k/2)}{\sqrt{35k/12}}\right)
\end{aligned}$$

$$\mathbf{P}\left(Z \geq \frac{99.5 - (7k/2)}{\sqrt{35k/12}}\right) > 0.05$$

In other words, we are looking for some quantile such that the area to the right exceeds 0.05. For this to be possible, the quantile must be *less* than $\Phi^{-1}(0.95)$. Let $c = \Phi^{-1}(0.95)$.

$$\frac{99.5 - (7k/2)}{\sqrt{35k/12}} < c$$

$$99.5 - \frac{7}{2}k < c\sqrt{\frac{35}{12}}\sqrt{k}$$

$$-\frac{7}{2}k - c\sqrt{\frac{35}{12}}\sqrt{k} + 99.5 < 0$$

Solving the equation, we get that k must be greater than 24.46 (approximately). Since k is discrete, we can round it up to 25. Therefore k should be greater than 25 so that $\mathbf{P}(S_k \geq 100) > 0.05$.

Question 4

Using MGFs, show that the sum of two independent Poisson random variables is a Poisson random variable. What is the parameter of the new random variable?

Let A be a general Poisson distributed random variable with parameter λ. We start by finding the MGF of A.

$$\begin{aligned}
M(t) &= \mathbf{E}\left(e^{tA}\right) \\
&= \sum_{a=0}^{\infty} e^{ta} \cdot p(a) \\
&= \sum_{a=0}^{\infty} e^{ta} e^{-\lambda} \frac{\lambda^a}{a!} \\
&= e^{-\lambda} \sum_{a=0}^{\infty} \frac{\left(\lambda e^t\right)^a}{a!} \\
&= e^{-\lambda} \cdot e^{\lambda e^t} \\
&= e^{\lambda(e^t - 1)}, \quad t \in \mathbb{R}
\end{aligned}$$

Now, let $X \sim \text{Poisson}(\lambda_1)$ and $Y \sim \text{Poisson}(\lambda_2)$, and $X \perp Y$. Then the MGF for $X + Y$ can be found as:

$$\begin{aligned}
\mathbf{E}\left(e^{t(X+Y)}\right) &= \mathbf{E}\left(e^{tX+tY}\right) \\
&= \mathbf{E}\left(e^{tX} \cdot e^{tY}\right) \\
&= \mathbf{E}\left(e^{tX}\right) \cdot \mathbf{E}\left(e^{tY}\right) && \text{(Since } X \perp Y\text{)} \\
&= e^{\lambda_1(e^t - 1)} \cdot e^{\lambda_2(e^t - 1)} && \text{(MGF of a Poisson rv)}
\end{aligned}$$

$$= e^{\lambda_1(e^t-1)+\lambda_2(e^t-1)}$$

$$= e^{(\lambda_1+\lambda_2)(e^t-1)}, \quad t \in \mathbb{R}$$

By the uniqueness of MGFs, it follows that

$$X + Y \sim \text{Poisson}(\lambda_1 + \lambda_2)$$

Question 5

Using the MGF of the exponential distribution, obtain all the moments of the exponential distribution.

To begin, let us consider the specific case of $X \sim \text{Exp}(\lambda = 1)$. Its MGF is

$$M(t) = \frac{1}{1-t}, \quad t < 1$$

We recognize the above as the formula for the sum of a geometric series. Working backwards, we have:

$$M(t) = \frac{1}{1-t} = \sum_{n=0}^{\infty} t^n = \sum_{n=0}^{\infty} n! \cdot \frac{t^n}{n!}, \quad |t| < 1 \tag{1}$$

It was shown in lecture that:

$$M(t) = \mathbf{E}\left(e^{tX}\right) = \mathbf{E}\left(\sum_{n=0}^{\infty} \frac{(tX)^n}{n!}\right) = \mathbf{E}\left(\sum_{n=0}^{\infty} \frac{t^n X^n}{n!}\right) = \sum_{n=0}^{\infty} \mathbf{E}\left(X^n\right) \frac{t^n}{n!}, \tag{2}$$

for t in some open interval about 0. Matching coefficients in (1) and (2), it follows that $\mathbf{E}(X^n) = n!$ for all $n \geq 0$.

By the properties of the exponential distribution, it is known that if $X \sim \text{Exp}(\lambda)$, for $k > 0$:

$$kX \sim \text{Exp}\left(\frac{\lambda}{k}\right)$$

Therefore, if $X \sim \text{Exp}(\lambda = 1)$, then:

$$\frac{1}{k} X \sim \text{Exp}(k)$$

Taking advantage of this, if $X \sim \text{Exp}(\lambda = 1)$, we define:

$$Y := \frac{X}{\lambda_Y} \sim \text{Exp}(\lambda_Y)$$

It follows that:

$$\mathbf{E}(Y^n) = \mathbf{E}\left(\frac{X^n}{\lambda_Y^n}\right) = \frac{\mathbf{E}(X^n)}{\lambda_Y^n} = \frac{n!}{\lambda_Y^n}$$

Question 6

Using the MGF of the standard normal distribution, obtain all the moments of the standard normal distribution.

From lecture, it was shown that the MGF of the standard normal is:

$$M(t) \;=\; e^{t^2/2}, \quad t \in \mathbb{R}$$

Performing a Taylor series expansion of the above, we have:

$$M(t) \;=\; e^{t^2/2} \;=\; \sum_{n=0}^{\infty} \frac{(t^2/2)^n}{n!} \;=\; \sum_{n=0}^{\infty} \frac{t^{2n}}{2^n\, n!} \;=\; \sum_{n=0}^{\infty} \frac{(2n)!}{2^n\, n!} \cdot \frac{t^{2n}}{(2n)!}$$

Matching coefficients once again using:

$$M(t) \;=\; \sum_{n=0}^{\infty} \mathbf{E}\left(X^n\right) \frac{t^n}{n!},$$

it follows that:

$$\mathbf{E}\left(Z^{2n}\right) \;=\; \frac{(2n)!}{2^n\, n!}$$

with the odd moments equal to zero due to the symmetry of the normal distribution.

It turns out that

$$\mathbf{E}\left(Z^{2n}\right) \;=\; \frac{(2n)!}{2^n\, n!} \;=\; (2n-1)!!$$

where $(2n-1)!!$ is the odd skip factorial, $n \geq 1$. For example:

$$\mathbf{E}\left(Z^2\right) = 1$$

$$\mathbf{E}\left(Z^4\right) = 3 * 1$$

$$\mathbf{E}\left(Z^6\right) = 5 * 3 * 1$$

Question 7

Let Z be a standard normal random variable. Compute the following probabilities:

(a) $\mathbf{P}(0 \le Z \le 2.17)$

(b) $\mathbf{P}(0 \le Z \le 1)$

(c) $\mathbf{P}(-2.50 \le Z \le 0)$

(d) $\mathbf{P}(-2.50 \le Z \le 2.50)$

(e) $\mathbf{P}(Z \le 1.37)$

(f) $\mathbf{P}(-1.75 \le Z)$

(g) $\mathbf{P}(-1.50 \le Z \le 2.00)$

(h) $\mathbf{P}(1.37 \le Z \le 2.50)$

(i) $\mathbf{P}(1.50 \le Z)$

(j) $\mathbf{P}(|Z| \le 2.50)$

Question 8

Consider the random sum

$$X = X_1 + X_2 + X_3 + \ldots + X_N$$

where

$$X_i \overset{\text{iid}}{\sim} \text{Exp}(\lambda) \quad \text{and} \quad N \sim \text{Geometric}(p), \quad X_i \perp N$$

Determine the distribution of X by finding its MGF.

$$\begin{aligned}
M(t) &= \mathbf{E}\left(e^{tX}\right) \\
&= \sum_{k=1}^{\infty} \mathbf{E}\left(e^{tX} \mid N = k\right) \cdot \mathbf{P}(N = k) && (*) \\
&= \sum_{k=1}^{\infty} \left(\frac{\lambda}{\lambda - t}\right)^k \cdot (1-p)^{k-1} \cdot p \\
&= \frac{\lambda p}{\lambda - t} \sum_{k=1}^{\infty} \left(\frac{\lambda}{\lambda - t}\right)^{k-1} (1-p)^{k-1} \\
&= \frac{\lambda p}{\lambda - t} \sum_{k-1=0}^{\infty} \left(\frac{\lambda}{\lambda - t} \cdot (1-p)\right)^{k-1} \\
&= \frac{\lambda p}{\lambda - t} \cdot \frac{1}{1 - \frac{\lambda - \lambda p}{\lambda - t}} && (\text{Require that } t < \lambda p) \\
&= \frac{\lambda p}{\lambda p - t}, \quad t < \lambda p
\end{aligned}$$

By the uniqueness of MGFs, it follows that $X \sim \text{Exp}(\lambda p)$. The property used in $(*)$ is known as the **law of total expectation**. This will be discussed later on in the course.