

# Problem Section 3 KEY

## Transformations of Random Variables

---

### Instructions

Please answer the following questions in the order in which they are posed. Add a few empty lines below each and write your answers there. Focus on answering in complete sentences. You will also need scratch paper/pen to work out the answers before typing it.

For help with formatting documents in RMarkdown, please consult R Markdown: The Definitive Guide. Another option is to search using Google.

---

### Exercises

1. Suppose  $X$  and  $Y$  are independent discrete random variables each with PMF

$$f(x) = \frac{1}{3} \quad x = 0, 1, 2.$$

Find the PMF of the random variable  $S = X + Y$ .

We have that  $S$  can take the following 5 values,  $S = 0, 1, 2, 3, 4$ .

We can think of each  $X$  and  $Y$  as 3 sided die, and  $S$  will be the sum. Since we have 2 3-sided die we have  $3 \times 3 = 9$  total combinations. Thus for each value of  $S$ , we can consider the fraction  $\frac{\text{number of ways to add to } S}{9}$ .

Thus we have the following PMF:

$s$	0	1	2	3	4
outcomes	(0,0)	(0,1); (1,0)	(0,2), (1,1), (2,0)	(1,2), (2,1)	(2,2)
$f(s)$	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{3}{9}$	$\frac{2}{9}$	$\frac{1}{9}$

Let's use the Theorem 16.1 to find the PDF of  $S$  so you can see how to use the formula:

$$\begin{aligned} f(s) &= P(S = s) \\ &= \sum_x f_1(x) f_2(s - x) \end{aligned}$$

where the possible values of  $S$  are  $s = 0, 1, 2, 3, 4$ .

For  $s \leq 2$ ,  $x$  takes values from  $0, 1, \dots, s$ . These  $x$  values give non-zero values for  $f_1(x)$  and also  $f_2(s - x)$ . In this case:

$$\begin{aligned} f(s) &= \sum_{x=0}^s f_1(x) f_2(s - x) \\ &= \sum_{x=0}^s \frac{1}{3} \times \frac{1}{3} = \frac{1}{9} \sum_{x=0}^s 1 \\ &= \frac{(s+1)}{9}. \quad s = 0, 1, 2 \end{aligned}$$

For  $s = 3$ , we have that  $x$  can take the value 1 or 2. These are the only values that make both  $f_1(x)$  and  $f_2(s - x)$  non-zero. In this case:

$$\begin{aligned} f(s) &= \sum_{x=1}^2 f_1(x)f_2(s-x) \\ &= \sum_{x=1}^2 \frac{1}{3} \times \frac{1}{3} = \frac{2}{9}, \quad s = 3 \end{aligned}$$

For  $s = 4$ , the value of  $x$  can only be 2 for both  $f_1(x)$  and  $f_2(s - x)$  to be non-zero. So

$$f(s) = f_1(2)f_2(2) = \frac{1}{9} \quad s = 4$$

This is an example where the formula is quite tedious to use. Better to just use the table of joint probabilities.

2. On an average weekday, about 67,000 cars cross the Evergreen Point (Hwy 520) floating bridge over Lake Washington. On weekends (Saturday and Sunday), the average is about 72,000 per day. Assume daily traffic counts are each approximately normal with given means and with standard deviation 1,000 cars/day, independent from day to day.
  - a. What is the distribution of the total number of cars crossing the bridge in a week (5 week days, 2 weekend days)?

Let  $X_1, \dots, X_5$  denote the number of cars on the 5 weekdays. More specifically we have that:

$$X_i \sim N(\mu = 67000, \sigma = 1000), i = 1, 2, 3, 4, 5$$

Similarly let  $Y_1, Y_2$  be the number of cars crossing on the weekends, so we have that:

$$Y_j \sim N(\mu = 72000, \sigma = 1000), j = 1, 2$$

We wish to find the distribution of  $Z = X_1 + X_2 + \dots + X_5 + Y_1 + Y_2 = \sum_{i=1}^5 X_i + \sum_{j=1}^2 Y_j$ .

Since these are independent random variables, by the properties of the normal distribution, we know  $Z$  will be normally distributed. We will have that:

$$\begin{aligned} E(Z) &= E\left(\sum_{i=1}^5 X_i + \sum_{j=1}^2 Y_j\right) \\ &= \sum_{i=1}^5 E[X_i] + \sum_{j=1}^2 E[Y_j] \text{ by linearity of expectation} \\ &= 5E[X_1] + 2E[Y_1] \text{ since each } X \text{ and } Y \text{ follows the same distribution} \\ &= 5 \times 67000 + 2 \times 72000 \\ &= 4.79 \times 10^5 \end{aligned}$$

Similarly we will have that:

$$\begin{aligned}
Var(Z) &= Var\left(\sum_{i=1}^5 X_i + \sum_{j=1}^2 Y_j\right) \\
&= \sum_{i=1}^5 Var[X_i] + \sum_{j=1}^2 Var[Y_j] \text{ by independence of the RV's} \\
&= 5Var[X_1] + 2Var[Y_1] \text{ since each X and Y follows the same distribution} \\
&= 5 \times 1000^2 + 2 \times 1000^2 \\
&= 7000000
\end{aligned}$$

Thus we have that  $SD(Z) = 2645.7513111$

Thus we have that the total number of cars crossing the bridge in a 7 day week, Z, has distribution:

$$Z \sim N(\mu = 4.79 \times 10^5, \sigma = \sqrt{7000000} = 2645.7513111)$$

- b. Calculate the probability that more than 482,000 cars cross the bridge in a week.

We wish to find  $P(Z > 482,000)$ . We can simply use pnorm to find this in R:

```
p_Z_482 <- 1 - pnorm(482000, mean = 5*67000 + 2*72000, sd = sqrt(7000000))
```

Thus we see that  $P(Z > 482,000) = 0.1284196$

3. Let  $q$  denote the 0.6 quantile (or 60th percentile) of a continuous distribution. What is the probability that the larger of two random variables drawn independently from this distribution will exceed  $q$ ?

Hint: Consider the complement

Denote continuous distribution Q, such that  $X_1, X_2 \sim Q$ . We have that  $q$  is the 60th percentile of Q. We wish to find  $P(\max(X_1, X_2) > q)$ . Using the hint we have that:

$$P(\max(X_1, X_2) > q) = 1 - P(\max(X_1, X_2) \leq q) = 1 - P(X_1 \leq q \cap X_2 \leq q)$$

Since  $X_1, X_2$  are independent RV's, we have that  $P(X_1 \leq q \cap X_2 \leq q) = P(X_1 \leq q) \times P(X_2 \leq q)$ .

Since  $q$  is the 60th percentile of Q, we know that  $P(X_1 \leq q) = P(X_2 \leq q) = .6$ . Thus we have that:

$$P(X_1 \leq q \cap X_2 \leq q) = P(X_1 \leq q) \times P(X_2 \leq q) = .6 \times .6 = .36$$

Thus putting it all together we have:

$$P(\max(X_1, X_2) > q) = 1 - P(\max(X_1, X_2) \leq q) = 1 - P(X_1 \leq q \cap X_2 \leq q) = 1 - .36 = .64$$

So the probability that the larger of the 2 RV's is greater than  $q$  is .64

4. Suppose a device has three independent components, all of whose lifetimes (in months) can be modeled by an exponential PDF with rate  $\lambda = 1$ .

How long should you expect to wait for the last item to fail?

Denote  $X_1, X_2, X_3$  denote the failure times of each component. We have that  $X_i \sim \exp(\lambda)$  for  $i=1,2,3$ .

We wish to find the expected time by which all three components fail. Since every component will have failed by the time the third component has failed we wish to find  $E[\max(X_1, X_2, X_3)]$ . Thus we can begin by finding the distribution of  $\max(X_1, X_2, X_3)$ , and taking its expected value.

We know by Theorem 16.6 that  $f_{max}(x) = n[F(x)]^{n-1}f(x)$ . For an exponential RV, we have that  $f(x) = \lambda \exp(-\lambda x)$  and that  $F(x) = 1 - e^{-\lambda x}$ . In this case we have that  $n=3$  and  $\lambda = 1$ .

Thus we have that:

$$f_{max}(x) = n[F(x)]^{n-1}f(x) = 3[1 - e^{-x}]^{3-1}e^{-x}$$

Doing some simplification we have that:

$$\begin{aligned} f_{max}(x) &= 3[1 - e^{-x}]^2 e^{-x} \\ &= 3e^{-x}[1 - 2e^{-x} + e^{-2x}] \\ &= 3e^{-x} - 6e^{-2x} + 3e^{-3x} \end{aligned}$$

Before we move onto the expectation let us remind ourselves about the expectation of exponential random variables. We have that if  $Y \sim \exp(\lambda^*)$  then  $E[Y] = \frac{1}{\lambda^*}$ . More specifically we have that:

$$\int_0^{\infty} y \lambda^* e^{-\lambda^* y} dy = \frac{1}{\lambda^*}$$

Using this fact we have that:

$$\begin{aligned} E(\max X_1, X_2, X_3) &= \int_0^{\infty} x f_{max}(x) dx \\ &= \int_0^{\infty} x [3e^{-x} - 6e^{-2x} + 3e^{-3x}] dx \\ &= 3 \int_0^{\infty} x e^{-x} - 3 \times 2 \int_0^{\infty} x e^{-2x} + 3 \int_0^{\infty} x e^{-3x} dx \\ &= 3 - 3 \frac{1}{2} + \frac{1}{3} \\ &= \frac{18}{6} - \frac{9}{6} + \frac{2}{6} \\ &= \frac{11}{6} \end{aligned}$$

Thus the expected time to replacement is  $\frac{11}{6}$  or 1.83 months. So while any component fails in 1 month on average, we can expect the maximum time to failure to be 1.83 months on average.