Problem Section 3

KEY

Monday Jan 22 2024

Learning Outcomes

The problems are designed to build conceptual understanding and problem-solving skills. The emphasis is on learning to find, evaluate and build confidence. The specific tasks include:

- Finding the distribution of the sum of two independent random variables
- Finding the mean and variance of the sum of independent normal random variables.
- Performing probability calculations with a normal distribution
- Be able to work out probabilities for the maximum of two continuous random variables
- Back up and support work with relevant explanations

Exercises

1. We can write the joint PMF of $\langle X, Y \rangle$ in a tabular form. Each (x, y) pair has probability $\frac{1}{9}$ since the random variables are independent.

	x		
y	0	1	2
0	1/9	1/9	1/9
1	1/9	1/9	1/9
2	1/9	1/9	1/9

We have that S can take the following 5 values, $S = 0, 1, 2, 3, 4$. Thus we have the following PMF for S:

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)
$f(s)$	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{3}{9}$	$\frac{2}{9}$	$\frac{1}{9}$

2.

a. Let $X_1, \ldots, X_5$ denote the number of cars crossing the bridge on the 5 weekdays. More specifically we are given that:

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

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

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

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

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

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

Similarly we will have that:

$$
\begin{aligned}
Var(T) &= Var(\sum_{i=1}^{5} X_i + \sum_{j=1}^{2} Y_j) \\
&= \sum_{i=1}^{5} Var[X_i] + \sum_{j=1}^{2} Var[Y_j] \text{ by independence of the RV's} \\
&= 5Var[X_1] + 2Var[Y_1] \text{ since each X and Y follows the same distribution} \\
&= 5 \times 1000^2 + 2 \times 1000^2 \\
&= 7000000
\end{aligned}
$$

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, W, has distribution:

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

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

```
p_t_482 <- 1 - pnorm(482000,mean = 5*67000 + 2*72000, sd = sqrt(7000000))
```

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

3.

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

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, 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.

a. Find the PDF of $X_{max} = \max\{X_1, X_2, X_3\}$

We know by Theorem 16.5 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:

$$f_{max}(x) = n[F(x)]^{n-1}f(x) = 3[1 - e^{-x}]^{3-1}e^{-x} \quad x \geq 0$$

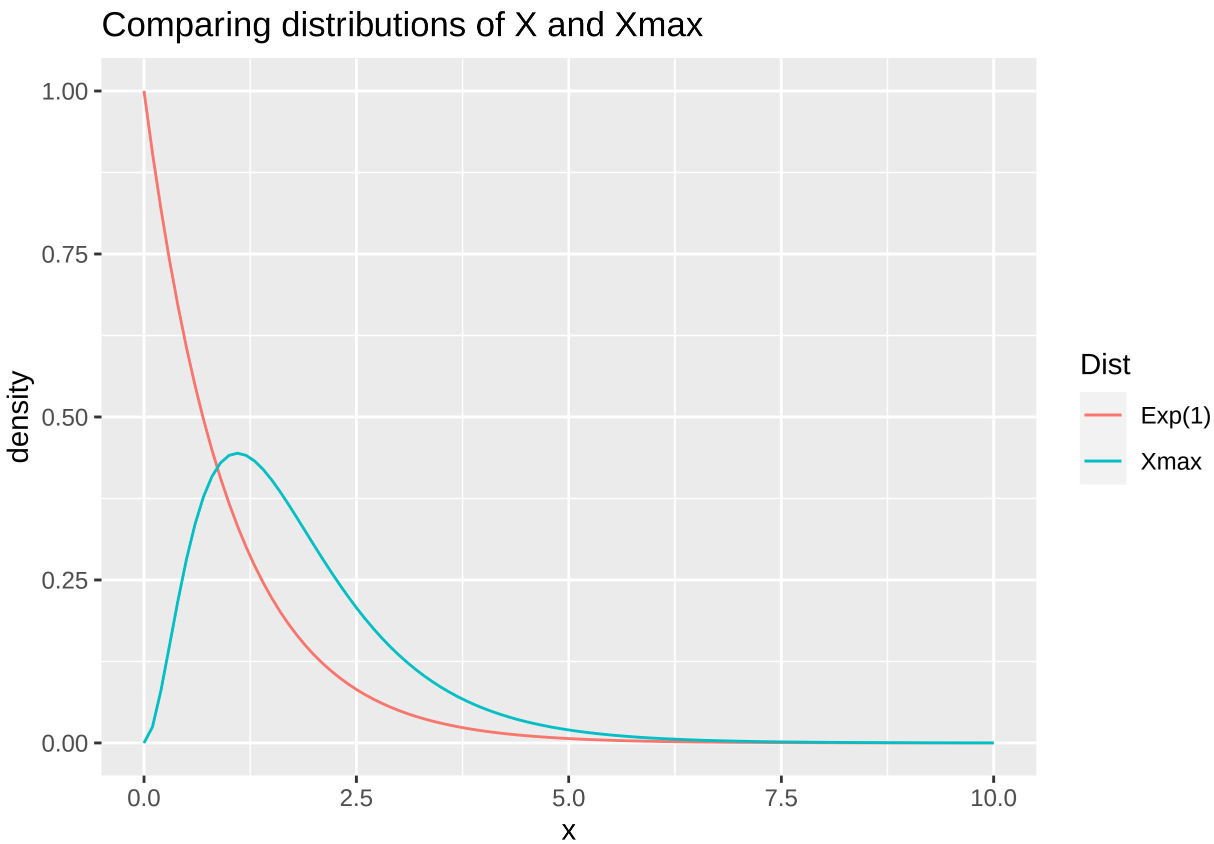
Doing some simplification we have that:

$$\begin{aligned} f_{max}(x) &= 3[1 - e^{-x}]^2 e^{-x} \\ &= 3e^{-x}[1 - 2e^{-x} + e^{-2x}] \\ &= 3e^{-x} - 6e^{-2x} + 3e^{-3x} \quad x \geq 0 \end{aligned}$$

b.

```r
pdf_max <- function(x){
  #return the PDF of Xmax at a given value x
  3*exp(-x)-6*exp(-2*x)+3*exp(-3*x)
}


ggplot(data = NULL) +
  #first layer is for X ~ Exp(1)
  geom_function(fun = dexp,
                args = list(rate = 1),
                mapping = aes(color = "Exp(1)" ),
                xlim=c(0,10) ) +
  #second layer is for PDF of X_{max}
  geom_function(fun = pdf_max,
                mapping = aes(color = "Xmax"),
                xlim=c(0,10)  ) +
  labs(x = "x",
       y = "density",
       title = "Comparing distributions of X and Xmax",
       color = "Dist")
```



Comparing distributions of X and Xmax
1.00
0.75
0.50
0.25
0.00
density
0.0
2.5
5.0
7.5
10.0
x
Dist
Exp(1)
Xmax