Homework 4 KEY

Combining 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 and show work whether we ask for it or not**. 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. (Floods) Suppose that many years of observation have confirmed that the annual maximum flood tide X (in feet) for a certain river can be modeled by the PDF

   $$f(x) = \frac{1}{20} \qquad 20 < x < 40$$

   The Army Corps of Engineers is planning to build a levee along a certain portion of the river and they want to make it high enough so there is only a 20% chance that the worst flood in the next thirty years will overflow the embankment. How high should the levee be?

   Be sure to define random variables, show your derivations and clearly highlight important intermediate steps so we don't miss them.

Define the RV's $X_1, \ldots, X_{30}$, that denote the flood tide height of the river during a given year. We have that for each X_i for $i = 1, \ldots, 30$ has a Uniform(20,40) distribution.

We wish to find the value k such that:

$$P(max(X_1, \ldots, X_{30}) \geq k) = .2$$

We have that:

$$F_{max}(x) = P(max(X_1, \ldots, X_{30}) \leq x) = (\frac{x-20}{20})^{30}$$

Thus we wish to find the value k such that:

$$P(max(X_1, \ldots, X_{30}) \geq k) = 1 - (\frac{k-20}{20})^{30} = .2$$

Solving for k we have:

$$
\begin{aligned}
1-(\frac{k-20}{20})^{30} &= .2 \\
(\frac{k-20}{20})^{30} &= .8 \\
\frac{k-20}{20} &= (.8)^{1/30} \\
k-20 &= 20(.8)^{1/30} \\
k &= 20(.8)^{1/30}+20 \\
&= 39.8517895
\end{aligned}
$$

Thus the Levee should be at least 39.8517895 feet tall.

2. (Two Poissons) Suppose $X_1, X_2, \ldots, X_n \overset{i.i.d.}{\sim} Pois(\lambda)$. That is, the PMF of each X is

$$f(x) = e^{-\lambda}\frac{\lambda^x}{x!}, \quad x = 0, 1, 2, \ldots$$

Show, using mathematical induction that

$$S = X_1 + X_2 + \cdots + X_n \sim Pois(n\lambda).$$

Hint: Look at Example 16.1 and Theorem 16.2 for the proof for the Binomial distribution.

We may prove this using induction.

Define our base case for n=2. We will show that if $X_1 \sim Pois(\lambda_1)$ independently of $X_2 \sim Pois(\lambda_2)$ then $X_1 + X_2 \sim Pois(\lambda_1 + \lambda_2)$. It makes sense to allow the rates to differ since we will need this in the induction step.

$$
\begin{aligned}
f(s) &= P(X_1 + X_2 = s) \quad s = 0, 1, 2, \ldots, \\
&= \sum_{x=0}^{s} f_{X_1}(x) f_{X_2}(s-x) \text{ using convolution} \\
&= \sum_{x=0}^{s} e^{-\lambda_1}\frac{\lambda_1^x}{x!}\, e^{-\lambda_2}\frac{\lambda_2^{s-x}}{(s-x)!} \\
&= \frac{e^{-(\lambda_1+\lambda_2)}}{s!}\sum_{x=0}^{s}\frac{s!}{x!(s-x)!}\lambda_1^x\lambda_2^{s-x} \\
&= \frac{e^{-(\lambda_1+\lambda_2)}}{s!}\sum_{x=0}^{s}\binom{s}{x}\lambda_1^x\lambda_2^{s-x} \\
&= \frac{e^{-(\lambda_1+\lambda_2)}}{s!}(\lambda_1+\lambda_2)^s \quad \text{binomial theorem}
\end{aligned}
$$

where the last sentence follows from the binomial expansion for $(\lambda_1 + \lambda_2)^s$.

Thus we have for n=2, $X_1 + X_2 \sim Pois(\lambda_1 + \lambda_2)$. In particular, if $\lambda_1 = \lambda_2 = \lambda$ then

$$X_1 + X_2 \sim Pois(2\lambda).$$

Now for our inductive step assume that for n=k, $S_k = X_1 + X_2 + \cdots + X_k \sim Pois(k\lambda)$. The induction hypothesis is that

$$S_{k+1} = X_1 + X_2 + \cdots + X_k + X_{k+1} \sim Pois((k+1)\lambda).$$

But this follows from the base case because $S_{k+1} \sim Pois(k\lambda)$ independently of $X_{k+1} \sim Pois(\lambda)$.

3. (Three normals) The random variables X_1, X_2 and X_3 are independent $Norm(\mu, 1)$. Let $\bar{X}$ be the (arithmetic) average of the three random variables. That is:

$$\bar{X} = \frac{1}{3}(X_1 + X_2 + X_3).$$

What is the probability that $Y = (X_1 - \bar{X}) > 1.6$?

Be sure to write the distribution of Y clearly, cite any results you use, then show the calculation including any code you write to do your calculations.

Hint Write Y as a linear combination of the X's. No further hints on this. You are expected to figure it out.

We have that:

$$Y = X_1 - \frac{1}{3}(X_1 + X_2 + X_3) = \frac{2}{3}X_1 - \frac{1}{3}X_2 - \frac{1}{3}X_3$$

Since X_1, X_2, X_3 are independent RV's, and Y is a linear combination of Normal RV's, Y is a normal RV.

We have from Theorem 15.1:

$$E[Y] = E[\frac{2}{3}X_1 - \frac{1}{3}X_2 - \frac{1}{3}X_3] = \frac{2}{3}E[X_1] - \frac{1}{3}E[X_2] - \frac{1}{3}E[X_3] = \frac{2}{3}\mu - \frac{1}{3}\mu - \frac{1}{3}\mu = 0.$$

We also have using the independence of the RV's that:

$$Var(Y) = Var(\frac{2}{3}X_1 - \frac{1}{3}X_2 - \frac{1}{3}X_3) = \frac{4}{9}Var(X_1) + \frac{1}{9}Var(X_2) + \frac{1}{9}Var(X_3) = \frac{4}{9} + \frac{2}{9} = \frac{6}{9}$$

Thus we have that:

$$Y \sim N(\mu = 0, \sigma = \frac{\sqrt{2}3}{)}$$

Thus using pnorm we have that $P(Y > 1.6) = 0.0250218$.

4. (Call center) Let the Poisson random variable X be the number of calls for technical assistance received by a computer company during the firm's nine normal workday hours. Suppose the average number of calls per hour is 7.0 and that each call costs the company \$50. Let Y be a Poisson random variable representing the number of calls for technical assistance received during a day's remaining fifteen hours. Suppose the average number of calls per hour is 4.0 for that time period and that each such call costs the company \$60.

   a. Let the random variable C denote the cost associated with the calls received during a 24 hour day. Write C as a function of X and Y.

   No hints on how to do this. You are expected to figure it out.

   We have that $C = 50X + 60V$

   b. Find the expected cost and the variance of the random variable C. Be sure to state any assumptions you make in order to do your calculation. And show your work.

Since we have that during normal workday hours, we have a rate of 7 calls an hour, we thus have 63 calls per 9 hours. Thus $X \sim Pois(\lambda = 63)$. Similarly in non-workday hours there is a rate of 60 calls per 15 hours. So $Y \sim Pois(\lambda = 60)$.

Using the fact that for a $Pois(\lambda)$ RV the expectation is λ, and linearity of expectation we have:

$$E[C] = E[50X + 60V] = 50E[X] + 60E[V] = 50 \times 63 + 60 \times 60 = \$6750$$

We also know that that Variance of a $Pois(\lambda)$ RV is equal to λ.

If we assume that the calls in workday and non-workday hours are independent we have using the properties of independent RV's for variance:

$$Var(C) = Var(50X + 60V) = 2500Var(X) + 3600Var(V) = 2500 \times 63 + 3600 \times 60 = 3.735 \times 10^5$$