Probability distributions
Lectures, questions and solutions

Nicky van Foreest and Ruben van Beesten

February 1, 2021

CONTENTS

1 LECTURE 1

Ex 1.1. Consider 12 football players on a football field. Eleven of them are players of FC Barcelone, the other one is an arbiter. We select a random player, uniform. This player must take a penalty. The probability that a player of Barcelone scores is 70%, for the arbiter it is 50%. Let $P \in \{A, B\}$ be r.v that corresponds to the selected player, and $S \in \{0,1\}$ be the score.

1. What is the PMF? In other words, determine $\mathsf{P}\{P = B, S = 1\}$ and so on for all possibilities.
2. What is $\mathsf{P}\{S = 1\}$? What is $\mathsf{P}\{P = B\}$?
3. Show that S and P are dependent.

An insurance company receives on a certain day two claims $X, Y \geq 0$. We will find the PMF of the loss $Z = X + Y$ under different assumptions.

The joint CDF $F_{X,Y}$ and joint PMF $p_{X,Y}$ are assumed known.

Ex 1.2. Why is it not interesting to consider the case $\{X = 0, Y = 0\}$?

Ex 1.3. Find an expression for the PMF of $Z = X + Y$.

Suppose $p_{X,Y}(i,j) = c\sum_{i,j} I_{i=j} I_{1\leq i\leq 4}$.

Ex 1.4. What is c?

Ex 1.5. What is $F_x(i)$? What is $F_Y(j)$?

Ex 1.6. Are X and Y dependent? If so, why, because $1 = F_{X,Y}(4,4) = F_X(4)F_Y(4)$?

Ex 1.7. What is $\mathsf{P}\{Z = k\}$?

Ex 1.8. What is $\mathsf{V}[Z]$?

Now take X, Y iid $\sim$ Unif($\{1,2,3,4\}$) (so now no longer $p_{X,Y}(i,j) \neq I_{i=j} I_{1\leq i\leq 4}$).

Ex 1.9. What is $\mathsf{P}\{Z = 4\}$?

Remark 1.1. We can make lots of variations on this theme.

1. Let $X \in \{1,2,3\}$ and $Y \in \{1,2,3,4\}$.
2. Take $X \sim$ Pois(λ) and $Y \sim$ Pois(μ). (Use the chicken-egg story)
3. We can make X and Y such that they are (both) continuous, i.e., have densities. The conceptual ideas[1] don't change much, except that the summations become integrals.
4. Why do people often/sometimes (?) model the claim sizes as iid $\sim$ Norm(μ, σ^2)? There is a slight problem with this model (can real claim sizes be negative?), but what is the way out?
5. The example is more versatile than you might think. Here is another interpretation.

   A supermarket has 5 packets of rice on the shelf. Two customers buy rice, with amounts X and Y. What is the probability of a lost sale, i.e., $\mathsf{P}\{X + Y > 5\}$? What is the expected amount lost, i.e., $\mathsf{E}[\max X + Y - 5, 0]$?

   Here is yet another. Two patients arrive in to the first aid of a hospital. They need X and Y amounts of service, and there is one doctor. When both patients arrive at 2 pm, what is the probability that the doctor has work in overtime (after 5 pm)n, i.e., $\mathsf{P}\{X + Y > 5 - 2\}$?

[1] Unless you start digging deeper. Then things change drastically, but we skip this technical stuff.

2 SOLUTIONS

s.1.1. Here is the joint PMF:

$$\mathsf{P}\{P = A, S = 1\} = \frac{1}{12}0.5 \qquad \mathsf{P}\{P = A, S = 0\} = \frac{1}{12}0.5 \tag{2.1}$$

$$\mathsf{P}\{P = B, S = 1\} = \frac{11}{12}0.7 \qquad \mathsf{P}\{P = B, S = 0\} = \frac{11}{12}0.3. \tag{2.2}$$

Now the marginal PMFs

$$\mathsf{P}\{S = 1\} = \mathsf{P}\{P = A, S = 1\} + \mathsf{P}\{P = B, S = 1\} = 0.042 + 0.64 = 0.683 = 1 - \mathsf{P}\{S = 0\}$$
$$\mathsf{P}\{P = B\} = \frac{11}{12} = 1 - \mathsf{P}\{P = A\}.$$

For independence we take the definition. In general, for all outcomes x, y we must have that $\mathsf{P}\{X = x, Y = y\} = \mathsf{P}\{X = x\}\mathsf{P}\{Y = y\}$. For our present example, let's check for a particular outcome:

$$\mathsf{P}\{P = B, S = 1\} = \frac{11}{12} \cdot 0.7 \neq \mathsf{P}\{P = B\}\mathsf{P}\{S = 1\} = \frac{11}{12} \cdot 0.683$$

The joint PMF is obviously not the same as the product of the marginals, which implies that P and S are not independent.

s.1.2. When the claim sizes are 0, then the insurance company does not receive a claim.

s.1.3. By the fundamental bridge,

$$\mathsf{P}\{Z = k\} = \sum_{i,j} I_{i+j=k} p_{X,Y}(i,j) \tag{2.3}$$
$$= \sum_{i,j} I_{i,j \geq 0} I_{j=k-i} p_{X,Y}(i,j) \tag{2.4}$$
$$= \sum_{i=0}^{k} p_{X,Y}(i, k-i). \tag{2.5}$$

s.1.4. $c = 1/4$ because there are just four possible values for i and j.

s.1.5. Use marginalization:

$$F_X(k) = F_{X,Y}(k, \infty) = \sum_{i \leq k} \sum_j p_{X,Y}(i,j) \tag{2.6}$$
$$= \frac{1}{4} \sum_{i \leq k} \sum_j I_{i=j} I_{1 \leq i \leq 4} \tag{2.7}$$
$$= \frac{1}{4} \sum_{i \leq k} I_{1 \leq i \leq 4} \tag{2.8}$$
$$= k/4, \tag{2.9}$$
$$F_Y(j) = j/4. \tag{2.10}$$

s.1.6. The equality in the equation must hold for all i, j, not only for $i = j = 4$. If you take $i = j = 1$, you'll see immediately that $F_{X,Y}(1,1) \neq F_X(1)F_Y(1)$:

$$\frac{1}{4} = F_{X,Y}(1,1) \neq F_X(1)F_Y(1) = \frac{1}{4}\frac{1}{4}. \tag{2.11}$$

s.1.7. $\mathsf{P}\{Z=2\} = \mathsf{P}\{X=1, Y=1\} = 1/4 = \mathsf{P}\{Z=4\}$, etc. $\mathsf{P}\{Z=k\} = 0$ for $k \notin \{2,4,6,8\}$.

s.1.8. Here is one approach

$$
\begin{aligned}
\mathsf{V}[Z] &= \mathsf{E}\left[Z^2\right] - (\mathsf{E}[Z])^2 && (2.12)\\
\mathsf{E}\left[Z^2\right] &= \mathsf{E}\left[(X+Y)^2\right] = \mathsf{E}\left[X^2\right] + 2\,\mathsf{E}[XY] + \mathsf{E}\left[Y^2\right] && (2.13)\\
(EZ)^2 &= (\mathsf{E}[X] + \mathsf{E}[Y])^2 && (2.14)\\
&= (\mathsf{E}[X])^2 + 2\,\mathsf{E}[X]\,\mathsf{E}[Y] + (\mathsf{E}[Y])^2 && (2.15)\\
&\Longrightarrow && (2.16)\\
\mathsf{V}[Z] &= \mathsf{E}\left[Z^2\right] - (\mathsf{E}[Z])^2 && (2.17)\\
&= \mathsf{V}[X] + \mathsf{V}[Y] + 2(\mathsf{E}[XY] - (\mathsf{E}[X]\,\mathsf{E}[Y])) && (2.18)\\
\mathsf{E}[XY] &= \sum_{ij} ij\, p_{X,Y}(i,j) = \frac{1}{4}(1+4+9+16) = \ldots && (2.19)\\
\mathsf{E}\left[X^2\right] &= \ldots && (2.20)
\end{aligned}
$$

The numbers are for you to compute.

s.1.9.

$$
\begin{aligned}
\mathsf{P}\{Z=4\} &= \sum_{i,j} I_{i+j=4}\, p_{X,Y}(i,j) && (2.21)\\
&= \sum_{i=1}^{4}\sum_{j=1}^{4} I_{j=4-i}\,\frac{1}{16} && (2.22)\\
&= \sum_{i=1}^{3} \frac{1}{16} && (2.23)\\
&= \frac{3}{16}. && (2.24)
\end{aligned}
$$