Probability distributions
Lectures, questions only

Nicky van Foreest and Ruben van Beesten

February 11, 2021

CONTENTS

1 LECTURE 1

Ex 1.1.

Ex 1.2. 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 1.3. Why is it not interesting to consider the case $\{X = 0, Y = 0\}$?

Ex 1.4. Find an expression for the PMF of $Z = X + Y$.

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

Ex 1.5. What is c?

Ex 1.6. What is $F_X(i)$? What is $F_Y(j)$?

Ex 1.7. Are X and Y dependent? If so, why, because $1 = F_{X,Y}(4,4) = F_X(4)F_Y(4)$?

Ex 1.8. What is $\mathsf{P}\{Z = k\}$?

Ex 1.9. What is $\mathsf{V}[Z]$?

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

Ex 1.10. 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 $X \sim \text{Pois}(\lambda)$ and $Y \sim \text{Pois}(\mu)$. (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 $\sim \text{Norm}(\mu, \sigma^2)$? There is a slight problem with this model (can real claim sizes be negative?), but what is the way out?

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

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), i.e., $\mathsf{P}\{X+Y>5-2\}$?

2 LECTURE 2

Read the problems of `memoryless\_excursions.pdf`. All the problems in that document relate to topics discussed in Sections BH.7.1 and BH.7.2, and quite a lot of topics you have seen in the previous course on probability theory.

3 LECTURE 3

Ex 3.1. We ask a married woman on the street her height X. What does this tell us about the height Y of her spouse? We suspect that taller/smaller people choose taller/smaller partners, so, given X, a simple estimator $\hat{Y}$ of Y is given by

$$\hat{Y} = aX + b.$$

(What is the sign of a if taller people tend to choose taller people as spouse?) But how to determine a and b? A common method is to find a and b such that the function

$$f(a,b) = \mathsf{E}\left[(Y - \hat{Y})^2\right]$$

is minimized. Show that the optimal values are such that

$$\hat{Y} = \mathsf{E}[Y] + \rho \frac{\sigma_Y}{\sigma_X}(X - \mathsf{E}[X]),$$

where ρ is the correlation between X and Y and where σ_X and σ_Y are the standard deviations of X and Y respectively.

Ex 3.2. Using scaling laws often can help to find errors. For instance, the prediction $\hat{Y}$ should not change whether we measure the height in meters or centimetres. In view of this, explain that

$$\hat{Y} = \mathsf{E}[Y] + \rho \frac{\mathsf{V}[Y]}{\sigma_X}(X - \mathsf{E}[X])$$

must be wrong.

Ex 3.3. N people throw their hat in a box. After shuffling, each of them takes out a hat at random. How many people do you expect to take out their own hat (i.e., the hat they put in the box); what is the variance? In BH.7.46 you have to solve this analytically. In the exercise here you have to write a simulator for compute the expectation and variance.

INDICATORS ARE GREAT functions, and I suspect you underestimated the importance of these functions. They help to keep your formulas clean. You can use them in computer code as logical conditions, or to help counting relevant events, something you need when numerically estimating multi-D integrals for machine learning for instance. And, even though I(=NvF) often prefer to use figures over algebra to understand something, when it comes to integration (and reversing the sequence of integration in multiple integrals) I find indicators easier to use.

Moreover, you should know that in fact, *expectation* is the fundamental concept in probability theory, and the *probability of an event is defined* as

$$\mathsf{P}\{A\} := \mathsf{E}[I_A]. \tag{3.1}$$

Thus, the fundamental bridge is actually an application of LOTUS to indicator functions. REREAD BH.4.4!

Here are some more examples

Ex 3.4. What is $\int_{-\infty}^{\infty} I_{0 \le x \le 3} \, \mathrm{d}x$?

Ex 3.5. What is

$$\int x \, I_{0 \le x \le 4} \, \mathrm{d}x? \tag{3.2}$$

When we do an integral over a 2D surface we can first integrate over the x and then over the y, or the other way around, whatever is the most convenient. (There are conditions about how to handle multi-D integral, but for this course these are irrelevant.)

Ex 3.6. What is

$$\iint xy \, I_{0 \le x \le 3} \, I_{0 \le y \le 4} \, \mathrm{d}x \, \mathrm{d}y? \tag{3.3}$$

Ex 3.7. What is

$$\iint I_{0 \le x \le 3} \, I_{0 \le y \le 4} \, I_{x \le y} \, \mathrm{d}x \, \mathrm{d}y? \tag{3.4}$$

4 LECTURE 4

Ex 4.1. BH.7.65 Let $(X_1,\ldots,X_k)$ be Multinomial with parameters n and $(p_1,\ldots,p_k)$. Use indicator r.v.s to show that $\mathrm{Cov}(X_i,X_j) = -np_ip_j$ for $i \neq j$.

Ex 4.2. Suppose (X,Y) are bivariate normal distributed with mean vector $\mu = (\mu_X,\mu_Y) = (0,0)$, standard deviations $\sigma_X = \sigma_Y = 1$ and correlation ρ_{XY} between X and Y. Specify the joint pdf of X and $X+Y$.

The following exercises will show how probability theory can be used in finance. We will look at the tradeoff between risk and return in a financial portfolio.

John is an investor who has \$10,000 to invest. There are three stocks he can choose from. The returns on investment (A,B,C) of these three stocks over the following year (in terms of percentages) follow a multinomial distribution. The expected returns on investment are $\mu_A = 7.5\%$, $\mu_B = 10\%$, $\mu_C = 20\%$. The corresponding standard deviations are $\sigma_A = 7\%$, $\sigma_B = 12\%$ and $\sigma_C = 17\%$. Note that risk (measured in standard deviation) increases with expected return. The correlation coefficients between the different returns are $\rho_{AB} = 0.7$, $\rho_{AC} = -0.8$, $\rho_{BC} = -0.3$.

Ex 4.3. Suppose the investor decides to invest \$2,000 in stock A, \$4,000 in stock B, \$2,000 in stock C and to put the remaining \$2,000 in a savings account with a zero interest rate. What the expected value of his portfolio after a year?

Ex 4.4. What is the standard deviation of the value of the portfolio in a year?

Ex 4.5. John does not like losing money. What is his probability of having made a net loss after a year?

John has a friend named Mary, who is a first-year EOR student. She has never invested money herself, but she is paying close attention during the course Probability Distributions. She tells her friend: “John, your investment plan does not make a lot of sense. You can easily get a higher expected return at a lower level of risk!”

Ex 4.6. Show that Mary is right. That is, make a portfolio with a higher expected return, but with a lower standard deviation.
*Hint: Make use of the **negative correlation** between C and the other two stocks!*

5 LECTURE 5

Ex 5.1. We have a continuous r.v. $X \geq 0$ with finite expectation. Use 2D integration and indicators to prove that

$$n\,\mathsf{E}[X] = \int_0^\infty x f(x)\,\mathrm{d}x = \int_0 G(x)\,\mathrm{d}x, \tag{5.1}$$

where $G(x)$ is the survival function.

Ex 5.2. Here is a nice geometrical explanation of the origin of the normal distribution. Suppose $z_0 = (x_0, y_0)$ is the target on a dart board at which Barney (our national darts hero) aims, but you can also interpret it as the true position of a star in the sky. Let z be the actual position at which the dart of Barney lands on the board, or the measured position of the star. For ease, take z_0 as the origin, i.e., $z_0 = (0,0)$. Then make the following assumptions:

1. The disturbance (x, y) has the same distribution in the x and the y direction.
2. The disturbance (x, y) along the x direction and the y direction are independent.
3. Large disturbances are less likely than small disturbances.

Show that the disturbance along the x-axis (hence y-axis) is normally distributed. You can use BH.8.19 as a source of inspiration. (This is perhaps a hard exercise, but the solution is easy to understand and very useful to memorize.)

Ex 5.3. We next find the normalizing constant of the normal distribution (and thereby offer an opportunity to practice with change of variables). For this purpose consider two circles in the plane: $C(N)$ with radius N and $C(\sqrt{2}N)$ with radius $\sqrt{2}N$. It is obvious that the square $S(N) = [-N, N] \times [N, N]$ contains the first circle, and is contained in the second. Therefore,

$$\iint_{C(N)} f_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y \leq \iint_{S(N)} f_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y \leq \iint_{C(\sqrt{2}N)} f_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y. \tag{5.2}$$

Now substitute the normal distribution in [5.2]. Then use polar coordinates to solve the integrals over the circles, and derive the normalization constant.

Ex 5.4. Let X, Y be iid with density f and support $[1, 10)$. Find an expression for the density of $Z = XY$. What is the (domain) support of Z? If $X, Y \sim \text{Unif}([1, 10))$, what is f_Z?

Remark 5.1. The above exercise is a step in the analysis of Benford's law that makes a statement on the first significant digit of numbers. Look it up on the web; it is a fascinating law. It's used to detect fraud by insurance companies and the tax department, but also to see whether the US elections in 2020 have been rigged, or whether authorities manipulate the statistics of the number of deceased by Covid. You can find the rest of the analysis in Section 5.5 of 'The art of probability for scientists and engineers' by R.W. Hamming.