Probability distributions EBP038A05
Lecture slides

Nicky van Foreest, Ruben van Beesten, Arpan Rijal

February 9, 2022

1

WEEK 1

1.1 LECTURE 1

Ex 1.1.1. Consider 12 football players on a football field. Eleven of them are players of F.C. Barcelona, the other one is an arbiter. We select a random player, uniform. This player must take a penalty. The probability that a player of Barcelona 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 1.1.2. Why is it not interesting to consider the case $\{X = 0, Y = 0\}$?

Ex 1.1.3. 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 1.1.4. What is c?

Ex 1.1.5. What is $F_X(i)$? What is $F_Y(j)$?

Ex 1.1.6. Are X and Y dependent? If so, why, because $1 = F_{X,Y}(4,4) = F_X(4)F_Y(4)$?

Ex 1.1.7. What is $\mathsf{P}\{Z = k\}$?

Ex 1.1.8. What is $\mathsf{V}[Z]$?

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

Ex 1.1.9. What is $\mathsf{P}\{Z = 4\}$?

Remark 1.1.10. 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.

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

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?

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

Ex 1.1.11. We have a continuous r.v. $X \geq 0$ with finite expectation. Use 2D integration and indicators to prove that

$$\mathsf{E}[X] = \int_0^\infty x f(x)\,\mathrm{d}x = \int_0 G(x)\,\mathrm{d}x, \tag{1.1.1}$$

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

Ex 1.1.12. A variation on BH.7.1. Alice is prepared to wait 20 minutes for Bob, while Bob doesn't want to wait longer than 10 minutes. What is the probability that they meet?

Use the fundamental bridge and indicator functions to write this probability as a 2D integral. Then use repeated integration to solve the 2D integral.

1.2 LECTURE 2

Ex 1.2.1. Let $L = \min\{X, Y\}$, where $X, Y \sim \text{Geo}(p)$ and independent. What is the domain of L? Then, use the fundamental bridge and 2D LOTUS to show that

$$\mathsf{P}\{L \geq i\} = q^{2i} \implies L \sim \text{Geo}(1-q^2).$$

Ex 1.2.2. Let $M = \max\{X, Y\}$, where $X, Y \sim \text{Geo}(p)$ and independent. Show that

$$\mathsf{P}\{M = k\} = 2pq^k(1-q^k) + p^2 q^{2k}.$$

Ex 1.2.3. Explain that

$$\mathsf{P}\{L = i, M = k\} = 2p^2 q^{i+k}\, I_{k>i} + p^2 q^{2i}\, I_{i=k}).$$

Ex 1.2.4. With the previous exercise, use marginalization to compute the marginal PMF $\mathsf{P}\{M=k\}$.x

Ex 1.2.5. Now take X, Y iid and $\sim \text{Exp}(\lambda)$. Use the fundamental bridge to show that for $u \leq v$, the joint CDF has the form

$$F_{L,M}(u,v) = \mathsf{P}\{L \leq u, M \leq v\} = 2\int_0^u (F_Y(v) - F_Y(x)) f_X(x)\,\mathrm{d}x.$$

Ex 1.2.6. Take partial derivatives to show that for the joint PDF,

$$f_{L,M}(u,v) = 2 f_X(u) f_Y(v)\, I_{u \leq v}.$$

2

WEEK 2

2.1 LECTURE 3

Ex 2.1.1. 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 2.1.2. 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 centimeters. 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 2.1.3. 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.

2.2 LECTURE 4

Ex 2.2.1. BH.7.65 Let $(X_1, \dots, X_k)$ be Multinomial with parameters n and $(p_1, \dots, p_k)$. Use indicator r.v.s to show that $\mathsf{Cov}\,[(] X_i, X_j) = -np_i p_j$ for $i \neq j$.

Ex 2.2.2. Suppose (X, Y) are bi-variate 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 trade off 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 2.2.3. 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 2.2.4. What is the standard deviation of the value of the portfolio in a year?

Ex 2.2.5. 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 2.2.6. 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!*

3

WEEK 3

3.1 LECTURE 5

HERE IS A NICE geometrical explanation of how the normal distribution originates.

Ex 3.1.1. 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 any 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.8.17 as a source of inspiration. (This is perhaps a hard exercise, but the solution is easy to understand and very useful to memorize.)

We next find the normalizing constant of the normal distribution (and thereby offer an opportunity to practice with change of variables).

Ex 3.1.2. 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,

$$\iint_{C(N)} f_{X,Y}(x, y)\,\mathrm{d}x\,\mathrm{d}y \le \iint_{S(N)} f_{X,Y}(x, y)\,\mathrm{d}x\,\mathrm{d}y \le \iint_{C(\sqrt{2}N)} f_{X,Y}(x, y)\,\mathrm{d}x\,\mathrm{d}y. \tag{3.1.1}$$

Now substitute the normal distribution of [3.1.1]. Then use polar coordinates (See BH.8.1.9) to solve the integrals over the circles, and derive the normalization constant.

BENFORD'S LAW 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. The next exercise is a first step in the analysis of Benford's law.

Ex 3.1.3. Let X, Y be iid with density f and support $[1, 10)$. Find an expression for the density of $Z = XY$. What is the support (domain) of Z? If $X, Y \sim \mathrm{Unif}([1, 10))$, what is f_Z?

3.2 LECTURE 6

Ex 3.2.1. Let $X \sim \text{Pois}(\lambda)$ and $Y \sim \text{Pois}(\mu)$ be independent. What is the distribution of $Z = X + Y$?

Ex 3.2.2. (BH.8.4.3.) Let $X_1, X_2, \ldots$ be i.i.d. $\text{Exp}(\lambda)$ distributed. Let $T_n = \sum_{k=1}^{n} X_k$. Show that T_n has the following pdf:

$$f_{T_n}(t) = \frac{\lambda^n}{(n-1)!} t^{n-1} e^{-\lambda t}, \quad t > 0. \tag{3.2.1}$$

That is, show that T_n follows a *Gamma distribution* with parameters n and λ. (We will learn about the Gamma distribution in BH.8.4.)

Ex 3.2.3. Let X, Y be i.i.d. $\mathcal{N}(0,1)$ distributed and define $Z = X + Y$. Show that $Z \sim \mathcal{N}(0,2)$ using a convolution integral.

4

WEEK 4

4.1 LECTURE 7

Ex 4.1.1. At the end of Story 2 of Bayes' billiards (BH.8.3.2) there is the expression

$$\beta(a,b) = \frac{(a-1)!(b-1)!}{(a+b-1)!}. \tag{4.1.1}$$

Derive this equation.

Ex 4.1.2. In the Beta-Binomial conjugacy story, BH take as prior $f(p) = I_{p\in[0,1]}$, and then they remark that when $f(p) \sim \beta(a,b)$ for general $a, b \in \mathbb{N}$, we must obtain the negative hypergeometric distribution. I found this pretty intriguing, so my question is: Relate the Bayes' billiards story to the story of the Negative Hypergeometric distribution, and, in passing, provide an interpretation of a and b in terms of white and black balls. Before trying to answer this question, look up the details of the negative hypergeometric distribution. (In other words, this exercise is meant to help you sort out the details of the remark of BH about the negative hypergeometric distribution.)

The next real exercise is about recursion applied to the negative hypergeometric distribution. But to get in the mood, here is short fun question on how to use recursion.

Ex 4.1.3. We have a chocolate bar consisting of n small squares. The bar is in any shape you like, square, rectangular, whatever. What is the number of times you have to break the bar such that you end up with the n single pieces?

Ex 4.1.4. Use recursion to find the expected number X of black balls drawn without replacement at random from an urn containing $w \geq 1$ white balls and b black balls before we draw 1 white ball. In other words, I ask to use recursion to compute $\mathsf{E}[X]$ for X a negative hypergeometric distribution with parameters $w, b, r = 1$ and show that

$$\mathsf{E}[X] = \frac{b}{w+1} \tag{4.1.2}$$

Ex 4.1.5. Extend the previous exercise to cope with the case $r \geq 2$. For this, write $N_r(w,b)$ for an urn with w white balls and b black balls, and r white balls to go.

4.2 LECTURE 8

Ex 4.2.1. Let X be a continuous random variable with a pdf

$$f_X(x) = \begin{cases} c, & \text{if } 0 \leq x \leq 4, \\ 0, & \text{otherwise.} \end{cases} \tag{4.2.1}$$

1. What is the value of c?

2. What is the distribution of X?

3. Do we need to know the value of c to determine the distribution of X?

Ex 4.2.2. Let X be a continuous random variable with a pdf

$$f_X(x) = c \cdot e^{-\frac{(x-4)^2}{8}}, \quad x \in \mathbb{R}. \tag{4.2.2}$$

1. What is the value of c?

2. What is the distribution of X?

3. Do we need to know the value of c to determine the distribution of X?

(BH.8.4.5) Fred is waiting at a bus stop. He knows that buses arrive according to a Poisson process with rate λ buses per hour. Fred does not know the value of λ, though. Fred quantifies his uncertainty of λ by the *prior distribution* $\lambda \sim \text{Gamma}(r_0, b_0)$, where r_0 and b_0 are known, positive constants with r_0 an integer.

Fred has waited for t hours at the bus stop. Let Y denote the number of buses that arrive during this time interval. Suppose that Fred has observed that $Y = y$.

Ex 4.2.3. Find Fred's (hybrid) joint distribution for Y and λ.

Ex 4.2.4. Find Fred's marginal distribution for Y. Use this to compute $\mathsf{E}[Y]$. Interpret the result.

Ex 4.2.5. Find Fred's posterior distribution for λ, i.e., his conditional distribution of λ given the data y.

Ex 4.2.6. Find Fred's posterior mean $\mathsf{E}[\lambda | Y = y]$ and variance $\mathsf{V}[\lambda | Y = y]$.

5

WEEK 5

5.1 LECTURE 9

Ex 5.1.1. The lifetime X of a machine is Exp(λ). Compute $\mathsf{E}[X \,|\, X \leq \tau]$ where τ is some positive constant.

Define $Y = I_{X \leq \tau}$. Use Adam's law and LOTP to show that $\mathsf{E}[X] = \mathsf{E}[\mathsf{E}[X \,|\, Y]] = 1/\lambda$. Observe that this is just a check on our results.

Ex 5.1.2. We have a station with two machines, one is working, the other is off. If the first fails, the other machine takes over. The repair time of the first machine is a constant τ. If the second machine fails before the first is repaired, the station stops working, i.e., is 'down'. Use a conditioning argument to find the expected time $\mathsf{E}[T]$ until the station is down when the lifetimes of both machines is iid $\sim$ Exp(λ).

Ex 5.1.3. We draw, with replacement, balls, numbered 1 to N, from an urn. Find a recursion to compute the expected number $\mathsf{E}[T]$ of draws necessary to see all balls.

Ex 5.1.4. Write code to compute $\mathsf{E}[T]$ for $N = 45$.

Ex 5.1.5. We draw, with replacement, balls, numbered 1 to N, from an urn, but 6 at a time (not just one as in the previous exercise). Find a recursion to compute the expected number $\mathsf{E}[T]$ of draws necessary to see all balls.

Ex 5.1.6. For the previous exercise, compute $\mathsf{E}[T]$ for $N = 45$.

5.2 LECTURE 10

We have a wooden stick of length 100 cm that we break twice. First, we break the stick at a random point that is uniformly distributed over the entire stick. We keep the left end of the stick. Then, we break the remaining stick again at a random point, uniformly distributed again, and we keep the left end again.

Ex 5.2.1. What is the expected length of the stick we end up with?

Ex 5.2.2. Now we change the story slightly. Every time we break a stick, we keep the *longest* part. What is the expected length of the remaining stick?

(Same story as last week; BH.8.4.5) Fred is waiting at a bus stop. He knows that buses arrive according to a Poisson process with rate λ buses per hour. Fred does not know the value of λ, though. Fred quantifies his uncertainty of λ by the *prior distribution* $\lambda \sim$ Gamma(r_0, b_0), where r_0 and b_0 are known, positive constants with r_0 an integer.

Fred has waited for t hours at the bus stop. Let Y denote the number of buses that arrive during this time interval. Suppose that Fred has observed that $Y = y$.

Ex 5.2.3. How many buses does Fred expect to observe? I.e., compute $\mathsf{E}[Y]$

(BH.9.3.10). An extremely widely used method for data analysis in statistics is *linear regression*. In its most basic form, the linear regression model uses a single explanatory variable X to predict a response variable Y. For instance, let X be the number of hours studied for an exam and let Y be the grade on the exam. The linear regression model assumes that the conditional expectation of Y is *linear* in X:

$$\mathsf{E}[Y|X] = a + bX. \tag{5.2.1}$$

Ex 5.2.4. Show that an equivalent way to express this is to write

$$Y = a + bX + \varepsilon, \tag{5.2.2}$$

where ε is a random variable (called the *error*) with $\mathsf{E}[\varepsilon|X] = 0$.

Ex 5.2.5. Solve for the constants a and b in terms of $\mathsf{E}[X]$, $\mathsf{E}[Y]$, $\mathsf{Cov}[X, Y]$, and $\mathsf{V}[Y]$.

WEEK 6

6

6.1 LECTURE 11

Ex 6.1.1. As a continuation of BH.9.2.4, we now ask you to use Eve's law to compute $\mathsf{V}[Y]$, the variance of the second break point.

Ex 6.1.2. A server (e.g., a machine, a mechanic, a doctor) spends a random amount T on a job. While the server works on the job, it sometimes gets interrupted to do other tasks $\{R_i\}$. Such tasks can be failures of the machine, and they need to be repaired before the machine can continue working again. (In the case of a mechanic, interruptions occur when the mechanic has to check some other machine or help other mechanics.) It is clear that such interruptions can only occur when the server is working. Assume that the interruptions arrive according to a Poisson process with rate λ, i.e., the number of failures $N|T \sim \text{Pois}(\lambda T)$. The interruptions $\{R_i\}$ are independent of T and form an iid sequence with common mean $\mathsf{E}[R]$ and variance $\mathsf{V}[R]$. (Typically, we get estimates for $\mathsf{E}[T]$ and $\mathsf{V}[T]$ from measurements.)

Use Adam and Eve's laws to express the expectation and variance of the total time S to complete a job in terms of $\mathsf{E}[T]$ and $\mathsf{V}[R]$.

This and the next exercise illustrate how to think about conditional expectation when the rv on which we condition is discrete.

Ex 6.1.3. We have a standard fair die and we paint red the sides with $1,2,3$ and blue the other sides. Let X be the color after a throw and Y the number. Use the definitions of conditional expectation, conditional variance and Adam's and Eve's law to compute $\mathsf{E}[X]$ and $\mathsf{V}[X]$. Compare that to the results you would obtain from using the standard methods; you should get the same result.

BTW, it's easy to make some variations on this exercise by painting other combinations of sides, e.g., only 1 and 2 are red.

Ex 6.1.4. We have a continuous r.v. X with PDF

$$f(x) = \begin{cases} 1/2, & 0 \le x < 1, \\ 1/4, & 1 \le x \le 3, \end{cases} \tag{6.1.1}$$

and 0 elsewhere. Compute $\mathsf{E}[X]$ and $\mathsf{V}[X]$.

Then define the r.v. Y such that $Y = I_{X<1} + 2\, I_{X\geq 1}$. Explain what information you obtain about X when somebody tells that $Y = 1$, or that $Y = 2$. Use conditioning on Y and Adam's and Eve's laws to recompute $\mathsf{E}[X]$ and $\mathsf{V}[X]$.

Again, with this exercise we can check whether we are applying all concepts in the correct way.

6.2 LECTURE 12

Some inequalities. Fill in either "≤" or "≥" at the location of the question mark.

Ex 6.2.1. $\mathsf{E}[XY]^2 \quad ? \quad \mathsf{E}\left[X^2\right]\mathsf{E}\left[Y^2\right]$.

Ex 6.2.2. $\mathsf{E}\left[\log(X)\right] \quad ? \quad \log(\mathsf{E}[X])$

Ex 6.2.3. $\mathsf{V}[Y] \quad ? \quad \mathsf{E}[\mathsf{V}[Y|X]]$

Ex 6.2.4. $\mathsf{E}[|X|] \quad ? \quad \sqrt{\mathsf{E}\left[X^2\right]}$

Ex 6.2.5. $\mathsf{P}\left\{X^2 \geq 4\right\} \quad ? \quad \mathsf{E}[|X|]/2$

Ex 6.2.6. Let $Z \sim N(0,1)$. Then, $\mathsf{P}\left\{Z > \sqrt{2}\right\} \quad ? \quad 1/e$.

We consider the height of a certain population of people (e.g., all students at the University of Groningen). For some reason, we don't know the value of the mean μ of the population, but we do know that the standard deviation σ is 10 cm. We use the sample mean $\bar{X}_n$ of an i.i.d. sample $X_1, \dots, X_n$, from the population (measured in cm) to estimate the true mean μ. We want to choose the sample size n in such a way that our estimate $\bar{X}_n$ is sufficiently reliable.

Ex 6.2.7. One measure of reliability of our estimator $\bar{X}_n$ is its standard deviation. Let's say we find our estimator reliable if its standard deviation is at most 1 cm. Give a lower bound for n for which we can guarantee that our esitmator is reliable in this sense.

Ex 6.2.8. Another measure of reliability of our estimator is given by the probability that our estimate is very bad. Specifically, we say our estimate is reliable if we can be 99% sure that our estimate is off by less than 5 cm. Give a lower bound for n for which we can guarantee that our estimator is reliable in this sense.

Little Mike needs to get ready for school. He must leave within 15 minutes, but there are two more things he needs to do: eat his breakfast and get dressed. The time it takes Mike to eat his breakfast has a mean value of 6 minutes with a standard deviation of 3 minutes. The time it takes Mike to get dressed has a mean value of 4 minutes with a standard deviation of 1 minute.

Ex 6.2.9. Give a lower bound for the probability that Mike will be ready for school in time.

7 WEEK 7

7.1 LECTURE 13

Ex 7.1.1. A shop receives demands with sizes $\{X_k\}$ for some product, for instance, cans with beans. Write $\mu = \mathsf{E}[X]$ and $\sigma^2 = \mathsf{V}[X]$.

Suppose n customers arrive. The shop has an initial inventory of I cans on the shelf. a) Use the CLT to determine I such that the probability that the total demand D remains below I is larger than some threshold α. b) The safety stock is defined as I minus the expected demand. How large is the safety stock? c) When is it reasonable to use the CLT to estimate I?

The next couple of exercises concentrate on the interpretation of expectation when dealing with real money, not just toy examples like throwing dice. We use this to explain why people pay for insurance, even though they have a negative expected value due to the payments to the insurance company. You should give some thought to exercises in the proper sequence.

Ex 7.1.2. Suppose we have one perfectly fair die. When the die lands 1, 2 you loose your investment, otherwise your investment gets doubled. You are given two options: bet all the money M you have and whatever you can lay your hands on (possibly including the total fortune of your parents, your friends, and so on), or not bet at all. Show that your expected gain is $M/3$. Given this, would you play this game?

Ex 7.1.3. You are given two options: bet any amount of money you like on the basis that with probability $p = 10^{-6}$ you win 1000001 times what you wagered, but with probability $1 - p$ you loose all. Would you bet all the money you have?

Ex 7.1.4. Consider the, so-called, St. Petersburg game in which we throw a fair coin until it comes up heads, and then we stop. When the coin lands heads on the nth throw, you receive 2^n Euros. Suppose your initial capital is m and you invest f, then your expected gain is $m - f + \sum_{n=1}^{\infty} 2^{-n} 2^n = \infty$. But, how much are you actually prepared to pay to enter this game? To resolve this, we don't focus on monetary expectation but instead on expected utility (see the previous exercise). Explain that, when investing f, the later is given by

$$\sum_{n=1}^{\infty} 2^{-n} \log(m - f + 2^n). \tag{7.1.1}$$

Ex 7.1.5. For $m = 200$, write a computer program to compute f' such that

$$\log m = \sum_{n=1}^{\infty} 2^{-n} \log(m - f' + 2^n). \tag{7.1.2}$$

What is the interpretation of this f'?

Ex 7.1.6. Suppose you pay a premium P to an insurance company to be protected against a random loss $L \in \{L_i, i \leq n\}$ occurring with probability $\mathsf{P}\{L = L_i\} = p_i$. Since insurance companies are very large, their utility is nearly linear in the range of L_i. Assume your utility function for money is the log function and you have an initial amount M of money. Explain that you and the insurance company are willing to do business when the premium P satisfies

$$\mathsf{E}[L] < P < M - \exp\left(\mathsf{E}\left[\log(M - L)\right]\right). \tag{7.1.3}$$

Ex 7.1.7. Use the Taylor series of $\log(1-x) \approx x + x^2/2$ and $e^x \approx 1 + x$ to see that when M quite a bit larger than $\mathsf{E}[L]$ that

$$M - \exp\left(\mathsf{E}\left[\log(M-L)\right]\right) \approx \mathsf{E}[L] + \frac{\mathsf{V}[L]}{2M} + \cdots \tag{7.1.4}$$

Interpret the result.

7.2 LECTURE 14

We shoot an arrow at a target. We aim at the center of the target. Our aim is not perfect though. We model our horizontal and vertical deviation from the target (in inches) by two independent standard normal random variables X and Y, respectively. (So $X < 0$ if we shoot to far to the left, for example.)

Ex 7.2.1. Compute the density function of the (Euclidean) distance from our arrow to the center of the target. Compare the result to the pdf of a standard normal random variable.

Ex 7.2.2. What is the expected distance from the center of the target?

Last week I stepped in dog poo two days in a row. This annoyed me and I decided that if the same happens to me again more than 5 times in the next 50 days, I will move to a neighbourhood with a lower dog population density.

Ex 7.2.3. Suppose that the probability I step in dog poo on a given day is 5%. Moreover, assume that the days are independent. Use a central limit theorem-based approximation to approximate the probability that I will decide to move as a result of the dog poo situation. (You may ignore the continuity correction.)

7.3 HINTS

h.1.1.11. Check the proof of BH.4.4.8

h.1.2.1. The fundamental bridge and 2D LOTUS have the general form

$$\mathsf{P}\left\{g(X,Y\right\} \in A\} = \mathsf{E}\left[I_{g(X,Y)\in A}\right] = \sum_i \sum_j I_{g(i,j)\in A}\,\mathsf{P}\left\{X=i, Y=j\right\}.$$

Take $g(i,j) = \min\{i,j\}$.

h.1.2.2. Use 2D LOTUS on $g(x,y) = I_{\max\{x,y\}=k}$.

h.3.2.1. Use the Binomial theorem:

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}, \tag{7.3.1}$$

for any non negative integer n.

h.3.2.2. Use mathematical induction.

h.4.1.4. For ease, write $N(w,b) = \mathsf{E}[X]$ for an urn with $w \geq r = 1$ white balls and b black balls. Then explain that

$$N(w,0) = 0 \quad \text{for all } w, \tag{7.3.2}$$

$$N(w,b) = \frac{b}{w+b}(1 + N(w,b-1)). \tag{7.3.3}$$

Then show that this implies that $N(w,b) = b/(w+1)$.

h.4.1.5. Explain that $N_0(w,b) = 0$ and

$$N_r(w,b) = \frac{w}{w+b} N_{r-1}(w-1,b) + \frac{b}{w+b}(1 + N_r(w,b-1)). \tag{7.3.4}$$

Then show that this implies that $N_r(w,b) = rb/(w+1)$.