Exercise Sheet 6 from Probability and Statistics — March 25–29, 2024

Recognizing Random Variables

1. The probability of a data breach at our server for each given day is 0.01, independently for each day. Let T be the number of days until the first data breach. What is the distribution of T, $\mathbb{E}(T)$, $\mathrm{var}(T)$? What is the probability that the server remains secure for an entire year?

2. Each software test can either find a bug (this we caount as success) or not (this we count as failure). Assume the probability of finding a bug in one test is 0.05, and a developer performs 20 independent tests, denote X as the number of bugs found. What is the distribution of X, $\mathbb{E}(X)$, $\mathrm{var}(X)$? What is the probability of finding exactly three bugs?

3. Historical data show that our server receives an average of 30 requests per minute. Use the Poisson distribution to determine the probability that the server will receive exactly 40 requests in the next minute.

Variance

Reminder:

- definition $\mathrm{var}(X) = \mathbb{E}((X - \mathbb{E}(X))^2)$
- theorem $\mathrm{var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$
- derived quantities: standard deviation $\sigma_X = \sqrt{\mathrm{var}(X)}$
- coefficient of variation $CV_X = \sigma_X/\mathbb{E}(X)$ (if $\mathbb{E}(X) > 0$)

4. Assume that solving one problem takes X minutes, where $X = 1, 2, \ldots,$ or 5. The duration is random (dependent on weather), and the probability function is $p_X(1) = p_X(2) = 0.1$, $p_X(3) = p_X(4) = 0.2$, $p_X(5) = 0.4$. From the past, we know that $\mathbb{E}(X) = 3.7$. Find $\mathrm{var}(X)$ and σ_X.

5. For random variables X, Y, we have $Y = aX + b$ (where a, b are real numbers).
(a) Express $\mathrm{var}(Y)$ in terms of $\mathrm{var}(X)$.
(b) Repeat for σ_X and CV_X.

6. Let $X \sim Bin(100, 0.5)$ and $Y \sim 10Bin(100, .05)$ (thus, $Y/10$ has a binomial distribution $Bin(100, .05)$). Compute (using formulas from the lecture) $\mathbb{E}(X)$, $\mathrm{var}(X)$, σ_X, CV_X and the same for Y.

Random Vectors

The joint probability function is defined by $p_{X,Y}(x, y) = P(X = x \,\&\, Y = y)$. *One-dimensional functions* p_X, p_Y are called marginal probability functions in this context. Recall how to find them from $p_{X,Y}$.

7. Let X, Y be the results of two independent rolls of a four-sided die (with numbers 1, …, 4).
(a) What is the probability function $Z_1 = \max(X, Y)$?
(b) What is the probability function $Z_2 = XY$?
[Hint: What values does the vector (X, Y) take if $\max(X, Y) = k$? Similarly, in the second part, if $XY = k$?]

$x \backslash y$	0	1	2
0	1/4	1/6	1/12
1	1/6	1/4	1/12

8. In this table you can find the joint probability function of random variables X, Y. These variables do not take any values other than those indicated.

(a) Find marginal distributions of X and Y. Compute $\mathbb{E}(X)$, $\mathbb{E}(Y)$.

(b) Are X and Y independent? In other words, does $p_{X,Y}(x, y) = p_X(x)p_Y(y)$ hold?

(c) Describe the distribution of $X+Y$ – i.e., find the probability function of the random variable $X+Y$. Compute $\mathbb{E}(X+Y)$ from here – verify if it equals $\mathbb{E}(X) + \mathbb{E}(Y)$.

(d) Describe the distribution of $X \cdot Y$. Compute $\mathbb{E}(XY)$ from here – verify if it equals $\mathbb{E}(X)\mathbb{E}(Y)$.

Conditional Expectation

$$\mathbb{E}(X \mid B) = \sum_{x \in Im(X)} x \cdot P(X = x \mid B)$$

$$\mathbb{E}(X) = \sum_i P(B_i) \cdot \mathbb{E}(X \mid B_i) \qquad B_1, B_2, \dots \text{ is a partition of } \Omega$$

9. In a TV quiz show, a participant can choose two questions. For question A, he estimates that he will correctly answer with a probability of 0.8 (and will earn 1,000 CZK for that). For question B, his success probability is only 0.5, but for a correct answer, he receives 2,000 CZK. The participant can choose the order in which to answer the questions in order to maximise his earnings. If he answers the first question incorrectly then the quiz terminates.

(a) What is the expected value of the winnings if he starts with question A?

(b) What if he starts with question B?

(c) Bonus: if the success probabilities are p_A, p_B, and the rewards m_A, m_B, how should the participant decide? * And what if there are more than two questions?

10. Flip a coin three times. Compute the expected number of heads given that the first head was in the second toss. In general, if we flip a coin n times, what is the expected number of heads given that the first head was in the kth toss.

More Practice Problems

11. Do the following *alternative definitions of variance* make sense? That is, do you learn anything interesting about random variable X when you find

(a) $\mathbb{E}(X - \mathbb{E}(X))$?

(b) $\mathbb{E}(|X - \mathbb{E}(X)|)$?

(c) $\mathbb{E}((X - \mathbb{E}(X))^k)$ for $k = 3, 4, \dots$.

12. Let $X \sim Poi(\lambda)$. Derive the relationships $\mathbb{E}(X) = \lambda$ and $\mathrm{var}(X) = \lambda$.

13. Let's consider a group of m married couples (i.e., a total of $2m$ individuals). Suppose that after ten years, each of these $2m$ people will still be alive with probability p, independently of the others. We do not consider possibilities of divorce, etc., so the couples are immutable.

Let L be the set of people who will be alive after ten years, and A their number (i.e., $A = |L|$). Furthermore, let B be the number of couples where both partners will be alive; thus, A, B are random variables satisfying $0 \le A \le 2m$ and $0 \le B \le m$. For each $a = 0, \dots, 2m$, we want to compute $\mathbb{E}(B \mid A = a)$.

(a) Let's consider a specific individual. What is the probability that they will be alive after ten years, given that $A = a$? In other words, if that person is x, what is $P(x \in L \mid A = a)$?

(b) Let's consider a specific married couple. What is the probability that both partners will be alive, given that $A = a$?

(c) Express B as the sum of m suitable indicator random variables.

(d) The linearity of expected value also holds for conditional expected value, i.e.,

$$\mathbb{E}(\sum_{i=1}^m X_i \mid J) = \sum_{i=1}^m \mathbb{E}(X_i \mid J),$$

for any event J and random variables $X_1, \ldots, X_m$. (You don't have to prove this.) Utilize this to compute $\mathbb{E}(B \mid A = a)$.

(e) What is the distribution of random variable A? (Either name it or write the probability function, i.e., determine $P(A = a)$.)

(f) For a chosen a-element set of people M, what is the probability that it exactly corresponds to the set of survivors? In other words, what is $P(L = M)$? And what about $P(L = M \mid A = a)$?

(g) For $m = 10$ and $a = 4$, verify the result by sampling in any programming language. If you use R, I recommend paying attention to the command **rbinom**(m,1,p) – it produces a vector with m numbers, each distributed according to $Bin(1, p)$, i.e., $Ber(p)$.

6th homework, due next week

14. We toss a coin three times. Let X be the number of heads in the first two tosses, and Y be the number of tails in the last two tosses.

(a) Determine the joint probability function $p_{X,Y}$ and also marginal probability functions p_X, p_Y.

(b) Are X and Y independent?

(c) Determine $P(X < Y)$.

(d) Determine the conditional probability function $p_{X|Y}$, i.e., the numbers $P(X = x \mid Y = y)$ for all values of x, y.