BM20A9301 Statistics – Exercise set 5

To be done by 5.–9.2.2024

Text in blue or red is not part of the problem or its solution. It's there as extra information to help you learn.

Exercise 1 (Covariance)**.** Two random variables, H and B have the joint distribution below, with marginal distributions already calculated. Context: flip two coins, H is the number of heads, B is how many tails before the first heads.

$P(H = h, B = b)$		b 0	1	2	$P(H = h)$
h	0	0	0	1/4	1/4
	1	1/4	1/4	0	2/4
	2	1/4	0	0	1/4
$P(B = b)$		2/4	1/4	1/4	

1. Calculate $E(H)$ and $E(B)$.
2. Create a table with 3 rows and 3 columns, of b and h like above, and fill it with the values of $(b - E(B)) \cdot (h - E(H)) \cdot P(H = h, B = b)))$.
3. Make another one with the values of $b \cdot h \cdot P(H = h, B = h)$.
4. Calculate $\mathrm{Cov}(H, B)$ in two ways, using both tables. Confirm that you get the same result. (Recall the formula $\mathrm{Cov}(X, Y) = E(X \cdot Y) - E(X) \cdot E(Y)$.)

Exercise 2 (Correlation & independence)**.** The joint distribution of two discrete r.v.'s X and Y is below:

		Y -1	0	1
	-1	0	1/6	1/6
X	0	1/3	0	0
	1	0	1/6	1/6

1. Determine the distribution, expected value and standard deviation of X.
2. Determine the distribution, expected value and standard deviation of Y.
3. Calculate the correlation of X and Y.
4. Figure out if X and Y are independent (make sure to explain why is that).

Exercise 3 (Average of dice rolls)**.** A fair dice is rolled several times in a row. The results are denoted $X_1, X_2, \ldots$ and they are independent. Denote the average of the first n rolls by $A_n = \frac{1}{n}(X_1 + \cdots + X_n)$.

1. Calculate the expected value and standard deviation of the r.v. X_1.

2. Determine the distribution of the r.v. $A_2 = \frac{1}{2}(X_1 + X_2)$.
   Hint: First determine the values A_2 can get. Then for small values of A_2 investigate which pairs (X_1, X_2) form that value. Then generalize.

3. Calculate the expected value and standard deviation of A_2. Compare the standard deviations of A_2 and X_1 by calculating their quotient.

4. Calculate the expected value and standard deviation of

$$A_{100} = \frac{1}{100}(X_1 + X_2 + \cdots + X_{100}).$$

   Compare the standard deviations of A_{100} and X_1 by calculating their quotient.

Exercise 4 (Normal distribution)**.** A shady battery manufacturer has no quality control so each battery is different. The batteries themselves are consistent, with some good and some bad. On average, the produced batteries last 50 hours before needing to be charged, with a standard deviation of 15 hours. The distribution is Gaussian. If you buy a battery from them, what is the probability that it will last between 50 and 70 hours without needing to be recharged?

Exercise 5 (Central limit theorem)**.** A city has reserved enough salt to keep the roads ice- and snowfree for up to 200cm snowfall in total. During a single winter day, it snows on average 4.5cm and the standard deviation is 2.5cm.

(a) Using the normal approximation, estimate the probability that there is enough salt for 50 winter days. **Hint:** You might need to make some assumptions to be able to make use of the central limit theorem.

(b) What extra assumptions did you have to make to solve (a)? Do you think those assumptions could be valid in this case?

Exercise 6 (Sample statistics)**.** Suppose that a dataset X consists of sampled data $17, 20, 18, 15, 18, 17, 15, 14, 16, 19$. Compute

(a) the mean,

(b) the variance, and

(c) the standard deviation.