STATISTICS FOR EES AND MEME—EXERCISE SHEET 4

1. A casino offers the following game: One dice is rolled, and if X is the number of pips, then the player wins $f(X) = 3 - (x-3)^2$ Euros, where negative values of $f(X)$ mean that the player loses money. Let $G = f(X)$ be the gain of a player in one game. Let T be the total gain of a player who played the game 100 times (where a negative gain is actually a loss).

(a) Calculate $\mathbb{E}G$ in two ways:

  (a) once directly with the definition of the expectation value $\sum_y y \cdot \Pr(G = y)$

  (b) once as $\sum_x f(x) \cdot \Pr(X = x)$

(b) Calculate the expectation value and the standard deviation of T.

(c) Which well-known distribution is can be used to approximate the distribution of T?

2.

(a) Imagine a board game in which you always roll a dice when it is your turn and move forward the number of steps corresponding to the number of dots shown on the dice. Calculate the expectation value and the standard deviation of the total number of steps you have moved forward after it has been your turn ten times (assuming that the board is large enough).

(b) Calculate this expectation value and standard deviation also for a variant of the game in which you can roll the dice once more when you had a six and move on accordingly. (But you cannot roll the dice a third time in one turn when you had two sixes.)

3. In a series of experiments with a male bird you played four times a recording of the song of a female of the same species and measured how long it took until the male replied with his song. The measured “reply times” in seconds were 1.52, 2.31, 1.24, 0.91.

(a) Calculate expectation value, variance and standard deviation of the from the data (as estimations of these parameters of the reply times of this male bird).

(b) Apply the transformation $g(x) = 1/x$ to the data to obtain the “reply rate” and calculate the expectation value, variance and standard deviation also for the transformed data.

(c) Apply another transformation f of the data, which consists in subtracting the minimum possible reaction time of 0.3 seconds and measuring the rest in milliseconds. Calculate expectation value, variance and standard deviation of the transformed data.

(d) To what extent could you also calculate expectation value, variance and standard deviation by applying the transformations g and f or modifications of them to the expectation value, variance and standard deviation calculated for the untransformed data?

4. In a published study you find a comparison between the genomes of two closely related species. In this study, 100 different alignable genomic regions of 1 kb each have been sampled, sequenced and compared. The average number of substitutions observed in a region was 7.8, where 4.2 and 3.6 were the average numbers of transitions and transversions, respectively. Of course, the actual numbers differed between the genomic regions, and the standard deviation was 2.4 for the transitions, 3.4 for the transversion and 4.9 for the total number of substitutions.
For your own research project the correlation of the number of transitions and the number of transversions in 1 kb regions is of interest. Can you calculate (more precisely, reasonably estimate) this correlation from the statistics given above or would you need to contact the authors of the study (or download the raw data)? If it is possible to calculate the correlation, what is the result? If not, why not?

5. Let $X_1, X_2, \ldots, X_n$ be independent random variables with the same distribution with finite variance σ^2. Proof that the corrected sample variance is unbiased, that is:

$$\mathbb{E}\left(\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \frac{1}{n}\sum_{j=1}^{n} X_j\right)^2\right) = \sigma^2.$$

6. Proof the following formulas or find counter examples:

(a) $\mathbb{E}(f(X)) = \sum_{x \in \mathcal{S}} f(x) \cdot \Pr(X = x)$

(b) $\text{Cov}(X, Y) = \mathbb{E}(X \cdot Y) - \mathbb{E}X \cdot \mathbb{E}Y$

(c) $\text{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}X)^2$

(d) $\text{Var}(X \cdot Y) = \text{Var}(X) \cdot \text{Var}(Y)$

(e) $\text{Cov}(a \cdot X, Y + Z) = a \cdot \text{Cov}(X, Y) + a \cdot \text{Cov}(X, Z)$