Probability Theory
2023/24 Period IIb
Instructor: Gilles Bonnet
Exam
21/6/2024
Duration: 2 hours

Solutions

1. (15 points) One deals a hand of 13 cards from a normal shuffled pack of 52 cards. What is the probability that it contains exactly one ace given that it contains exactly 2 kings?
Remark: Give your answer in the form of a *fraction*, where the numerator and denominator might be product of binomial coefficients.

Solution: Let A be the event of exactly one ace, and KK the event of exactly two kings. Then

$$\mathbb{P}(A|KK) = \frac{\mathbb{P}(A \cap KK)}{\mathbb{P}(KK)},$$

where

$$\mathbb{P}(A \cap KK) = \frac{\binom{4}{1}\binom{4}{2}\binom{44}{10}}{\binom{52}{13}}, \text{ and } \mathbb{P}(KK) = \frac{\binom{4}{2}\binom{48}{11}}{\binom{52}{13}}.$$

Thus

$$\mathbb{P}(A|KK) = \frac{\binom{4}{1}\binom{4}{2}\binom{44}{10}}{\binom{4}{2}\binom{48}{11}} = \frac{\binom{4}{1}\binom{44}{10}}{\binom{48}{11}}.$$

This was not required, but you can simplify the fraction to get the following:

$$\mathbb{P}(A|KK) = \frac{4! \cdot 44! \cdot 11! \cdot 37!}{1! \cdot 3! \cdot 10! \cdot 34! \cdot 48!} = \frac{4 \cdot 11 \cdot 37 \cdot 36 \cdot 35}{48 \cdot 47 \cdot 46 \cdot 45} = \frac{2^4 \cdot 3^2 \cdot 5 \cdot 7 \cdot 11 \cdot 37}{2^5 \cdot 3^3 \cdot 5 \cdot 23 \cdot 47} = \frac{7 \cdot 11 \cdot 37}{2 \cdot 3 \cdot 23 \cdot 47} = 0.44...$$

2. Let X be a random variable with expectation μ_X and variance σ_X^2, and Y a random variable with expectation μ_Y and variance σ_Y^2. Let $\rho_{X,Y}$ be the correlation between X and Y. Express the following quantities in terms of μ_X, μ_Y, σ_X, σ_Y, and $\rho_{X,Y}$.

   (a) (10 points) $\mathrm{Var}(X - 3Y)$;

   Solution:

$$\begin{aligned}\mathrm{Var}(X - 3Y) &= \mathrm{Cov}(X - 3Y, X - 3Y) \\ &= \mathrm{Cov}(X, X) + 9 \cdot \mathrm{Cov}(Y, Y) - 2 \cdot 3 \cdot \mathrm{Cov}(X, Y) \\ &= \sigma_X^2 + 9\sigma_Y^2 - 6\rho_{X,Y}\sigma_X\sigma_Y.\end{aligned}$$

(b) (10 points) $\mathbb{E}((3X-5)(2Y+1))$.

Solution: Note that

$$\mathrm{Cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) \quad \Longrightarrow \quad \mathbb{E}(XY) = \rho_{X,Y}\sigma_X\sigma_Y + \mu_X\mu_Y.$$

Hence,

$$\begin{aligned}\mathbb{E}((3X-5)(2Y+1)) &= \mathbb{E}(6XY + 3X - 10Y - 5)\\ &= 6\mathbb{E}(XY) + 3\mathbb{E}(X) - 10\mathbb{E}(Y) - 5\\ &= 6(\rho_{X,Y}\sigma_X\sigma_Y - \mu_X\mu_Y) + 3\mu_X - 10\mu_Y - 5.\end{aligned}$$

3. Let $N \in \mathbb{N}$ and $p \in (0,1)$.

A gambler has a starting capital of i€, with $i \in \mathbb{N}$. He repeatedly plays a game where he bets 1€ each time. With probability p he wins and he receives his money back plus an additional 1€, and with probability $1-p$ he loses his 1€.

If his capital reaches 0€ he is of course no longer able to play. And, if he manages to gather a capital of N€, he will go home contently.

Let p_i denote the probability that the gambler goes home contently, given that he starts with i€.

(a) (4 points) Show that for $1 \le i \le N-1$ we have the following recursive formula for p_i:

$$p_i = pp_{i+1} + (1-p)p_{i-1}.$$

(b) (4 points) For $1 \le i \le N-1$, express p_{i+1} in terms of p and p_1.

(c) (6 points) For $0 \le i \le N$, express p_i in terms of i and N.

Solution: This is Section 14.1.2 of the lecture notes.

4. Let $S = \sum_{r=1}^{n} X_r$ be a sum of independent Bernoulli random variables X_r taking values in $\{0,1\}$, where $\mathbb{E}(X_r) = p_r$ and $\mathbb{E}(S) = \mu > 0$.

(a) (5 points) Show that

$$\mathbb{E}(e^{tX_r}) \le \exp[p_r(e^t - 1)], \quad t \in \mathbb{R},\ r \in \{1, \dots, n\}.$$

Solution: We have

$$\mathbb{E}[e^{tX_r}] = (1-p_r)e^0 + p_r e^t = 1 + p_r(e^t - 1) \le \exp[p_r(e^t - 1)].$$

(b) (5 points) Show that

$$\mathbb{E}(e^{tS}) \le \exp[\mu(e^t - 1)], \quad t \in \mathbb{R}.$$

Solution: We have

$$\mathbb{E}(e^{tS}) = \mathbb{E}(e^{t\sum_{r=1}^{n} X_r}) = \mathbb{E}\left(\prod_{r=1}^{n} e^{tX_r}\right) = \prod_{r=1}^{n} \mathbb{E}(e^{tX_r}),$$

where the last equality holds because of the independence of the random variables X_r's. Therefore, with the first question of the exercise, we get

$$\mathbb{E}(e^{tS}) \le \prod_{r=1}^{n} \exp[p_r(e^t - 1)] = \exp\left[\sum_{r=1}^{n} p_r(e^t - 1)\right] = \exp\left[\mu(e^t - 1)\right].$$

(c) (10 points) Show that

$$\mathbb{P}(S > (1+\epsilon)\mu) \le \exp\left(-\mu\left[(1+\epsilon)\log(1+\epsilon) - \epsilon\right]\right), \quad \epsilon > 0.$$

Hint: You might want to use Markov's inequality in combination with the bound established in the previous question, and then choose t appropriately.

Solution: We can now use Markov's inequality to get

$$\begin{aligned}
\mathbb{P}(S > (1+\epsilon)\mu) &= \mathbb{P}(e^{tS} > e^{t(1+\epsilon)\mu}) \\
&\le \frac{\mathbb{E}(e^{tS})}{e^{t(1+\epsilon)\mu}} \\
&\le \exp\left[\mu(e^t - 1) - t(1+\epsilon)\mu\right] \\
&= \exp\left[\mu(e^t - 1 - t(1+\epsilon))\right].
\end{aligned}$$

Choosing $t = \log(1+\epsilon)$, we get

$$\begin{aligned}
\mathbb{P}(S > (1+\epsilon)\mu) &\le \exp\left[\mu\left((1+\epsilon) - 1 - \log(1+\epsilon)(1+\epsilon)\right)\right] \\
&= \exp\left[\mu\left(\epsilon - \log(1+\epsilon)(1+\epsilon)\right)\right] \\
&= \exp\left[-\mu\left((1+\epsilon)\log(1+\epsilon) - \epsilon\right)\right].
\end{aligned}$$

5. There has been an election for the student council and there were only two candidates, let us call them A and B. The voters have cast their ballots into a sealed box. Suppose that there were a votes for A and b votes for B.

   The committee in charge of counting the votes takes out the ballots one by one. The order in which they take out the ballots is completely random, because they have shaken the box really well and they don't look inside while they take the ballots out.

   (a) (6 points) Show that

$$\mathbb{P}(\text{ the last ballot they take out of the box is a vote for } A\text{ }) = \frac{a}{a+b}.$$

Solution: All permutations of the $a+b$ votes are equally likely. The number of ways to get a vote for A last is $a\cdot(a+b-1)!$. (Choose the last vote to be one of the a votes for A and arrange the other $a+b-1$ votes in a sequence arbitrarily.) Hence

$$\mathbb{P}(\text{last ballot is vote for } A) = \frac{a\cdot(a+b-1)!}{(a+b)!} = \frac{a}{a+b}.$$

The committee have a blackboard at their disposal, which they've divided into two parts, the left half and the right half, separated by a vertical line. If there is a vote for A the committee make a "tally" mark on the left board and if there is a vote for B they make a "tally" mark on the right side.

(Tallies on a black board.)

We are interested in the probability that at all times during the counting process there are more tallies on the left than on the right side of the board. That is, the probability that, for all $1 \leq i \leq a+b$, among the first i votes there are strictly more in favour of A than in favour of B.

(b) (15 points) Show that

$$\mathbb{P}(A \text{ is ahead during the entire counting process}) = \begin{cases} \frac{a-b}{a+b} & \text{if } a \geq b, \\ 0 & \text{otherwise.} \end{cases}$$

(Hint: Use induction on a **and** *b with induction hypothesis "for all (a',b') with $a' \leq a, b' \leq b$ and $(a,b) \neq (a',b')$ the statement holds", and use part (a).)*

Solution: *Base case.* The statement is clearly true when $a,b \leq 1$.

Induction step. Let a,b be arbitrary nonnegative integers, and assume the statement is true for all a',b' with $a' \leq a, b' \leq b, (a,b) \neq (a',b')$. If $a=0$ or $b=0$ then clearly the formula holds, so we may assume $a,b \geq 1$. Also, if $a \leq b$ then it is clearly impossible for A to be always ahead. So then the probability is zero and hence the formula is correct. We can assume $a > b \geq 1$ from now on.

We condition on the *last* vote. Let $E = \{\text{last vote is for } A\}$. Note that the probability that A was ahead at all times *given that E occurred* is the same as the probability that A is always ahead in a situation with $a-1$ votes for A and b votes for B. Hence by the induction hypothesis

$$\mathbb{P}(A \text{ always ahead}|E) = \frac{a-b-1}{a+b-1}.$$

Similarly, the probability that A was ahead at all times *given that E did not occur* is the same as the probability that A is always ahead in the situation with a votes for A and $b-1$ votes for B. So, again using the induction hypothesis:

$$\mathbb{P}(A \text{ always ahead}|E^c) = \frac{a-b+1}{a+b-1}.$$

Using the first part of the exercise, it follows that

$$
\begin{aligned}
\mathbb{P}(A \text{ always ahead}) &= \mathbb{P}(E)\mathbb{P}(A \text{ always ahead}|E) + \mathbb{P}(E^c)\mathbb{P}(A \text{ always ahead}|E^c) \\
&= \frac{a}{a+b} \cdot \frac{a-b-1}{a+b-1} + \frac{b}{a+b} \cdot \frac{a-b+1}{a+b-1} \\
&= \frac{a^2 - ab - a + ab - b^2 + b}{(a+b)(a+b-1)} \\
&= \frac{(a-b)(a+b-1)}{(a+b)(a+b-1)} = \frac{a-b}{a+b}.
\end{aligned}
$$