# Resit exam probability probability theory (WIKR-06)

9 July 2020, 8.30 – 12.00

- Work on this exam if the **5th digit** of your student id is **even**.
- Before the start of the exam, everybody taking part in the exam must sign the student declaration in the exam environment.
- To check for possible fraud, an unannounced sample of students will be contacted soon after the exam.
- The answers need to be written by hand, scanned and submitted within the time limit. You must upload your exam in a single pdf file.
- Every exercise needs to be handed in on a separate sheet.
- Write your name and student number **on every sheet**.
- It is forbidden to communicate with other persons during the exam, except with the course instructor.
- The only tools and aids that you are allowed to use are a non-programmable calculator (not a phone!), and the following material from the nestor course environment:
  a) The pdf file of the lecture notes (not videos, not scribbles).
  b) The pdf files of the tutorial problems.
  c) The pdf files of the homework problems.
  d) The pdf files of the solutions to the homework problems.
- Always give a short proof of your answer or a calculation to justify it, or clearly state the facts from the lecture notes or homework you are using.
- Simplify your final answers as much as possible.
- **NOTA BENE.** Using separate sheets for the different exercises, solving the exam corresponding to your student number, writing your name and student number on all sheets, and submitting all sheets in a single pdf is worth 10 out of the 100 points.

**Problem 1 (a:4, b:6, c:8, d:4, e:8 pts).**
The joint pdf of a random vector $(X,Y)$ equals

$$f_{X,Y}(x,y) := e^{-c|x-2y|-c|x+2y|}.$$

a) Determine $c$.

b) Determine the joint pdf $f_{U,V}(u,v)$ of $U := 2X+4Y$ and $V := 2X-4Y$. Are $U$, $V$ independent?

c) Determine the marginal pdf's $f_X(x)$ and $f_Y(y)$. Are $X$, $Y$ independent? Are $X,Y$ uncorrelated?

d) Determine and simplify $f_{X|Y}(x|y)$. The final result must not contain an expression of the form $|\cdot|$.

e) Let $Z$ be a random variable with pdf $f_Z(z) := f_{X|Y}(z|6)$. Compute $\mathbb{P}(|Z| \leqslant 1 \,|\, Z \in [-12,12])$ and $\mathbb{P}(Z > 200 \,|\, Z > 100)$.

**Note.** If you could not solve part d), then you may assume (incorrectly) for part e) that $f_{X|Y}(x|y) = \frac{1}{12y} 1_{[-3y,3y]}(x) + \frac{1}{2} e^{-(x-3y)} 1_{[3y,\infty)}(x)$.

**Problem 2 (a:6, b:6, c:6, d:2, e:4, f:6 pts).**
A factory produces light bulbs with independent random life times $\{\Delta_n\}_{n\geqslant 1}$. The life time $\Delta_n$ of the $n$th light bulb is a random variable with pdf $f(x) = 2e^{-2x}$.

a) Show that $\mathbb{P}(E) \geqslant \limsup_{n\to\infty} \mathbb{P}\big(\sum_{i\leqslant n}(\Delta_i - 1/2) \geqslant \sqrt{n}\big)$, where

$$E := \Big\{ \sum_{i\leqslant n}(\Delta_i - 1/2) \geqslant \sqrt{n} \text{ for infinitely many } n \geqslant 1 \Big\},$$

and conclude that $\mathbb{P}(E) > 0$.

Let $P$ be a Poisson random variable with parameter $\lambda = 6$ and set $N := \#\{n \geqslant 1 : \Delta_1 + \cdots + \Delta_n \leqslant 3\}$.

b) Show that $\mathbb{P}(P = 0) = \mathbb{P}(N = 0)$ and $\mathbb{P}(P = 1) = \mathbb{P}(N = 1)$.

c) Let $Y$ be a random variable on $\{0, 1, \dots\}$ with pmf $f_Y(k) = \frac{3}{4}4^{-k}$, and assume that $P$ and $Y$ are independent. Show that the pmf of $P+Y$ satisfies

$$f_{P+Y}(k) = 3e^{18}4^{-k-1}\mathbb{P}(Z \leqslant k),$$

where $Z$ is a Poisson random variable with parameter 24.

Set $U_i := 4e^{-2\Delta_i}$.

d) Show that $U_i$ is uniformly distributed on $[0, 4]$.

e) For $x \geqslant 0$, let $\lfloor x \rfloor$ be the largest integer smaller than $x$. Compute the probability that $\lfloor U_1 \rfloor$, $\lfloor U_2 \rfloor$ and $\lfloor U_3 \rfloor$ are pairwise distinct.

f) Show that the random variables $\big\{\sqrt[n]{U_1 \cdots U_n}\big\}_{n\geqslant 1}$ converge in probability and determine the limit.

**Problem 3 (a:5, b:3, c:8, d:4, e:2, f:8 pts).**
Sara the sorceress can perform a magic trick turning snails into beads. There are blue and red snails, as well as blue and red beads. Sara has a big bag containing 250 blue and 750 red snails. A blue snail turns into a blue bead with probability 90% and into a red bead with probability 10%. A red snail turns into a red bead with probability 95% and into a blue bead with probability 5%.

a) Sara draws a snail at random from her bag and performs her trick. What is the probability to craft a blue bead? Assume that the crafted bead is indeed blue. What is the probability that a crafted bead emerged from a blue snail?

Sara can undo her trick, reverting the bead into the snail it originated from and putting it back into her bag. Hence, the composition of snails in the bag remains the same. Sara now repeatedly performs and undoes her trick.

b) Let $N$ be the number of trials until crafting a blue bead for the first time. Compute the mgf $M_N(t)$ of $N$.

c) Suppose that the first crafted bead is red. Compute the expected number of trials until crafting two blue beads in a row.

We now subdivide the 1,000 snails at random into 200 groups consisting of 5 snails each. Independent of its color, each snail has a random number of speckles, which is uniformly distributed on $\{1,\dots,500\}$.

d) What is the probability that in at least one of the 200 groups all snails have the same number of speckles?

e) From each of the 200 groups, one representative is selected at random. What is the probability that all 200 representatives have a different number of speckles? You do not need to compute the numerical value of your result.

Finally, assume that among the 1,000 snails there are four *special* ones having precisely 23 speckles.

f) What is the probability that all special snails end up in the same group? How does this answer change if we additionally know that no group contains precisely one special snail?

**Solutions**
**a)**

$$\int_{\mathbb{R}}\int_{\mathbb{R}} e^{-c|x-2y|-c|x+2y|}\mathrm{d}x\mathrm{d}y = \int_{\mathbb{R}}\int_{\mathbb{R}} e^{-c|x|-c|x+4y|}\mathrm{d}x\mathrm{d}y = \frac{1}{2c}\int_{\mathbb{R}} e^{-c|x|}\mathrm{d}x = \frac{1}{c^2}.$$

Hence, $c = 1$.

**b)** Consider the transformation $g(x,y) = (2x-4y, 2x+4y)$. Then, the Jacobian is 16. By the transformation formula, the joint pdf of $U, V$ equals

$$f_{U,V}(u,v) = \frac{1}{16}e^{-\frac{c}{2}(|u|+|v|)}.$$

$U$ and $V$ are independent because the pdf factorizes.
**c)** By symmetry, we may assume $x > 0$.

$$f_X(x) = \int_{\mathbb{R}} e^{-c|x-2y|-c|x+2y|}\mathrm{d}y = \int_0^\infty e^{-c|x-y|-c|x+y|}\mathrm{d}y = xe^{-2cx} + \int_x^\infty e^{-2cy}\mathrm{d}y = (x+\frac{1}{2c})e^{-2cx}.$$

Similarly,

$$f_Y(y) = \int_{\mathbb{R}} e^{-c|x-2y|-c|x+2y|}\mathrm{d}x = 4\int_0^\infty e^{-c|2x-2y|-c|2x+2y|}\mathrm{d}x = 4ye^{-4cy} + 4\int_y^\infty e^{-4cx}\mathrm{d}x = (4y+\frac{1}{c})e^{-4cy}.$$

$X$ and $Y$ are not independent because $f_X(x)f_Y(y) \neq f_{X,Y}(x,y)$. Now, $X = (U+V)/4$ and $Y = (U-V)/8$, so that $\mathsf{Cov}(X,Y) = 2^{-5}(\mathsf{Var}(V) - \mathsf{Var}(U)) = 0$.

**d)** By symmetry, we may assume $x, y > 0$. Then,

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \begin{cases} \frac{c}{4cy+1} & \text{if } x \leqslant 2y, \\ \frac{ce^{-2c(x-2y)}}{4cy+1} & \text{if } x > 2y. \end{cases}$$

**e)** By part d), the random variable $Z$ is uniformly distributed on $[-12,12]$ when conditioned on the event $\{Z \in [-12,12]\}$. Hence, $\mathbb{P}(|Z| \leqslant 1 \,|\, Z \in [-12,12]) = 1/12$. Furthermore, $Z - 100$ is exponentially distributed with parameter $2c$ when conditioned on the event that $Z > 100$. Thus, $\mathbb{P}(Z > 200 \,|\, Z > 100) = e^{-200c}$.

**Solutions**

**a)** Note that

$$E = \bigcap_{m \geqslant 1} \bigcup_{n \geqslant m} E_n,$$

where $E_n = \sum_{i \leqslant n}(\Delta_i - 1/2) \geqslant \sqrt{n}$. Then,

$$\mathbb{P}(E) = \lim_{m\to\infty} \mathbb{P}(\cup_{n\geqslant m} E_n) \geqslant \lim_{m\to\infty} \sup_{n \geqslant m} \mathbb{P}(E_n) = \limsup_{n\to\infty} \mathbb{P}(E_n).$$

By the CLT, $n^{-1/2}\sum_{i\leqslant n}(\Delta_i - 1/2)$ converges in distribution to a normal random variable, so that $\mathbb{P}(E_n)$ converges to a non-zero probability.

**b)**

$$\mathbb{P}(N = 0) = \mathbb{P}(\Delta_1 > 3) = e^{-6} = \mathbb{P}(P = 0),$$

and

$$\mathbb{P}(N = 1) = \mathbb{P}(\Delta_1 \leqslant 3, \Delta_1 + \Delta_2 \geqslant 3) = \int_0^3 2e^{-2x}\mathbb{P}(\Delta_2 > 3 - x)\mathrm{d}x = \int_0^3 2e^{-6}\mathrm{d}x = \mathbb{P}(P = 1).$$

**c)**

$$f_{P+Y}(k) = \sum_{i \leqslant k} \mathbb{P}(Y = i)\mathbb{P}(P = k - i) = e^{-\lambda} \sum_{i\leqslant k} \frac{3}{4} 4^{-i} \frac{\lambda^{k-i}}{(k-i)!} = 3e^{-\lambda}4^{-k-1} \sum_{i \leqslant k} \frac{(4\lambda)^{k-i}}{(k-i)!} = 3e^{3\lambda}4^{-k-1}\mathbb{P}(Z \leqslant k).$$

**d)**

$$\mathbb{P}(U_i \leqslant r) = \mathbb{P}(2\Delta_i \geqslant -\log(r/4)) = r/4.$$

**e)** The probability that all are pairwise distinct equals $4 \cdot 3! \cdot 4^{-3} = 3/8$. We first choose which of the 4 intervals is not covered and then have 3! possible orderings.

**f)** We claim that the limit is $4/\sqrt{e}$. After taking logarithms, we reduce the claim to showing that $\frac{1}{n}\sum_{i\leqslant n} 2\Delta_i$ converges to 1/2 in probability. But this is just the LLN.

**Solutions**

**a)** Let $S, B \in \{\mathsf{b}, \mathsf{r}\}$ denote the colors of the snail and bead, respectively. Then, $\mathbb{P}(S = \mathsf{b}) = 0.25$, so that

$$\mathbb{P}(B = \mathsf{b}) = \mathbb{P}(S = \mathsf{b})\mathbb{P}(B = \mathsf{b} \,|\, S = \mathsf{b}) + \mathbb{P}(S = \mathsf{r})\mathbb{P}(B = \mathsf{b} \,|\, S = \mathsf{r}) = 0.25 \cdot 0.9 + 0.75 \cdot 0.05 = 0.2625.$$

Moreover, by Bayes,

$$\mathbb{P}(S = \mathsf{b} \,|\, B = \mathsf{b}) = \frac{\mathbb{P}(S = \mathsf{b}, B = \mathsf{b})}{\mathbb{P}(B = \mathsf{b})} = \frac{0.25 \cdot 0.9}{0.2625} = 0.857143.$$

**b)** The random variable $N$ is geometrically distributed with parameter $p = 0.2625$. Thus,

$$M_N(t) = p \sum_{k \geqslant 1} (1-p)^{k-1} e^{tk} = \frac{pe^t}{1 - (1-p)e^t}.$$

**c)** First, the expected number of additional trials until seeing the first blue bead is $1/p$. We call such a configuration a *run*, and note that together with the first step, the expected number of trials in a run is $1 + 1/p$. Now, letting $R$ denote the number of runs before seeing two blue beads for the first time, we note that also $R$ is geometrically distributed with parameter $p$. Hence, the expected number of trials until seeing the first time two blue beads in a row equals

$$1+\mathbb{E}[R](1+1/p) = 1+1/p+1/p^2.$$

**d)** The probability that all snails in one group have the same number of speckles equals $500^{-4}$. Hence, the probability that there exists at least one such group is given by $1-(1-500^{-4})^{200}$.

**e)** $500 \cdot 499 \cdots 301/500^{200}$.

**f)** Enumerate the snails at random from 1 to 1,000 and assume that the groups are formed according to this numbering. Then, there are 200 possibilities to put the 4 special snails into a group and 5 possibilities to order them inside the group. Hence, the probability equals

$$\frac{200 \cdot 5}{\binom{1000}{4}}.$$

For the second question, there is just one more option, namely two groups with two special snails each. There are $\binom{200}{2}$ possibilities to choose the two groups and $\binom{5}{2}^2$ to choose the positions of the special snails within the groups. Hence, we arrive at the conditional probability

$$\frac{200 \cdot 5}{200 \cdot 5 + \binom{200}{2}\binom{5}{2}^2}.$$