Probability Theory: Solutions for Resit 2020

These are the solutions to the resit of 2020. Note that this was an online exam of 3.5 hours and given by a different lecturer than now so it need not be representative for this years exam.

I think the solutions provided by in the other document are lacking, so here are some extra remarks. Moreover note that this was a *very* hard exam because this professor seemingly did not believe in resits, so it is highly unlikely your final will be of similar difficulty. If you are able to do these exercises you will probably be fine for the final.

Problem 1

---

a

As is standard for this type of problem, you use that a PDF has to integrate to 1. and check for which values of c this works out. So

$$1 = \iint_{\mathbb{R}^2} f_{X,Y} \, \mathrm{d}A = \iint_{\mathbb{R}^2} e^{-c|x-2y|-c|x+2y|} \, \mathrm{d}x \, \mathrm{d}y$$

Perform the change of coordinates

$$x \mapsto x - 2y, \qquad y \mapsto y;$$

which has jabobian

$$\det \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} = 1,$$

So that

$$\begin{aligned} 1 &= \iint_{\mathbb{R}^2} e^{-c|x|} e^{-c|x+4y|} \, \mathrm{d}x \, \mathrm{d}y, \\ &= \int_{\mathbb{R}} e^{-c|x|} \int_{\mathbb{R}} e^{-c|x+4y|} \, \mathrm{d}y \, \mathrm{d}x. \end{aligned}$$

Note that

$$\begin{aligned} \int_{\mathbb{R}} e^{-c|x+4y|} \, \mathrm{d}y &= \left[\int_{-\infty}^{-x/4} + \int_{-x/4}^{\infty} \right] e^{-c|x+4y|} \, \mathrm{d}y \\ &= e^{cx} \int_{-\infty}^{-x/4} e^{4cy} \, \mathrm{d}y + e^{-cx} \int_{-x/4}^{\infty} e^{4cy} \, \mathrm{d}y \\ &= \frac{1}{2c} \end{aligned}$$

So the whole thing evaluates to

$$1 = \frac{1}{2c} \int_{\mathbb{R}} e^{-c|x|} \, \mathrm{d}x = \frac{1}{c} \int_0^{\infty} e^{-cx} \, \mathrm{d} = \frac{1}{c^2}.$$

In the steps we assumed the integrals are convergent, which is not the case if $c = -1$, hence only $c = 1$ is a solution.

b

As is standard, we will write a function $g : \mathbb{R}^2 \to \mathbb{R}^2$ such that $f \circ g$ is the joint pdf. So

$$g : (x, y) \mapsto (2x - 4y, 2x + 4y)$$

suffices. This function has jabobian

$$\det g' = \begin{vmatrix} 2 & -4 \\ 2 & 4 \end{vmatrix} = 16$$

which we have to divide out to keep the integral equal to 1, giving

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

This can be factored as

$$\frac{1}{16} e^{-|u|/2} e^{-|v|/2}$$

which are PDFs, so the joint PDF factors and hence variables are independent.

c

To obtain f_X we integrate the joint PDF over all possible values of Y. Since this is even as a function of x we assume $x > 0$ without loss of generality. Note that the integrand is also even in y, hence we can just consider the positive part. The solution manual also implicitly performs the change of variables $y \mapsto 2y$ to ignore this factor 2. Showing all steps we get

$$\begin{aligned}
f_X &= \int_{\mathbb{R}} e^{-c|x-2y|-c|x+2y|} \, \mathrm{d}, \\
&= 2 \int_0^\infty e^{-|x-2y|-|x+2y|} \, \mathrm{d}y, \\
&= 2e^{-x} \int_0^\infty e^{-|x-2y|} e^{-2y} \, \mathrm{d}y, \\
&= 2e^{-x} \left[\int_0^{x/2} + \int_{x/2}^\infty \right] e^{-|x-2y|} e^{-2y} \, \mathrm{d}y, \\
&= 2e^{-x} \left[\int_0^{x/2} e^{-(x-2y)} e^{-2y} \, \mathrm{d}y + \int_{x/2}^\infty e^{x-2y} e^{-2y} \, \mathrm{d}y \right] \\
&= 2e^{x} \left[e^{-x} x/2 + e^{x} \int_{x/2}^\infty e^{-4y} \, \mathrm{d}y \right], \\
&= xe^{-2x} + \frac{1}{2} e^{-2x}.
\end{aligned}$$

We get f_Y by swapping x and $2y$ around, since this labeling is arbitrary. The random variables are not independent, because $f_X \cdot f_Y \neq f_{X,Y}$. They are uncorrelated because U, V are independent, so

$\mathrm{cov}(U, V) = 0$ hence

$$\begin{aligned}
\mathrm{cov}\,(X, Y) &= \mathrm{cov}\left(\frac{U+V}{4}, \frac{U-V}{8}\right) \\
&= \frac{1}{32}\,\mathrm{cov}(U, U) - \frac{1}{32}\,\mathrm{cov}(V, V) \\
&= \frac{1}{32}\left[\mathrm{var}(U) - \mathrm{var}(V)\right]
\end{aligned}$$

Note that U and V have identical PDFs so the variances are equal and the covariance is 0.

d

Using the definition of conditional probability

$$\begin{aligned}
f_{X|Y}(x|y) &= \frac{f_{X,Y}}{f_Y} \\
&= \frac{e^{-|x-2y|-|x+2y|}}{(4y+1)e^{-4y}} \\
&= \begin{cases} \dfrac{e^{x+2y-x-2y}}{(4y+1)e^{-4y}} = \dfrac{1}{4y+1}, & x \leq 2y, \\ \dfrac{e^{-x-2y-x-2y}}{(1+4y)e^{-4y}} = \dfrac{e^{-2x-4y}}{1+4y}, & \text{else.} \end{cases}
\end{aligned}$$

e

Note that the first probability has a uniform distribution, so the answer is just the length of the interval where $|z| \leq 1/12$ so $1/12$. For the other part, recognise an exponential distribution to get e^{-200}.

Problem 2

a

Trick: write E as a limit of sets which are nicely behaved.

$$E_n = \bigcap_{m\geq 1}\bigcup_{n\geq m}\left\{n : \sum_{i=1}^{n}\Delta_i - 1/2 \geq \sqrt{n}\right\}$$

thus

$$\mathbb{P}(E) = \mathbb{P}\left(\bigcup_{n\geq m} E_m\right)$$

Now the result follows by the fact that the supremum can never be larger than the union of all the elements which are in any of the E_n. The remainder of the question follows immediately from the central limit theorem.

b

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

and for the second one split the interval up

$$\mathbb{P}(N=1) = \mathbb{P}(\Delta_1 \leq 3, \Delta_2 \geq 3-\Delta_1) = \int_0^3 2e^{-3x}\mathbb{P}(\Delta_2 > 3-x)\,\mathrm{d}x = \mathbb{P}(P=1)$$

c

Discrete variable, so use a sum. Sum over all values of Y and P that work.

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

d

If the CDF is a constant then that corresponds with the CDF of a uniform distribution. And indeed

$$\begin{aligned}\mathbb{P}(U_i \leq x) &= \mathbb{P}(\ln U_i/4 \leq \ln x/4),\\ &= \mathbb{P}\left(\Delta_i \geq -\frac{1}{2}\ln(x/4)\right) = \int_{-\ln(x/4)/2}^{\infty} 2e^{-2y}\,\mathrm{d}y,\\ &= -e^{-2y}\Big|_{-\ln(x/4)/2}^{\infty} = x/4.\end{aligned}$$

e

Combinatronic exercise, each of the U_i is uniform on $[0,4]$, so $\lfloor U_i \rfloor$ takes 4 values, namely 0 on $[0,1)$, 1 on $[1,2)$, 2 on $[2,3)$, and 3 on $[3,4)$. The point 4 can be ignored because the probability of a single point is 0 in a continuous distribution.

So we need 3 of the U_i to land in different intervals. Pick the first U_i, it can land in any interval. Now for the second, there are 3 intervals where it can land, so the probability of it being distinct is $3/4$ (uniform distribution), and lastly there are $2/4 = 1/2$ intervals where the last can land. Since the Δ_i are independent we can multiply all this together

$$\frac{3}{4}\frac{1}{2} = \frac{3}{8}.$$

f

Each of the U_i have distribution $4e^{2\Delta_i}$, so heuristically we would expect

$$\sqrt[n]{U_1 \cdot U_n} \sim \left(4^n e^{-\sum_{i=1}^n 2\Delta_i}\right)^{1/n} \sim 4\left(e^{-n}\right)^{1/n} = 4/\sqrt{e}$$

and indeed

$$\mathbb{P}\left((U_1 \cdots U_n)^{1/n} = x\right) = \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^n \ln U_i = \ln x\right) = \mathbb{P}\left(\frac{4}{n}\sum_{i=1}^n \Delta_i = y\right)$$

recall the set E from part a and conclude by the weak law of large numbers that this is convergent.

Problem 3

a

Either snail can be turned into a blue bead, simply multiply the probabilities of drawing a certain snail with that of it turning blue, so

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

For the second part, Bayes' theorem tells us

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

b

The distribution is geometric with $p = 0.2625$. The MGF is given as

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

c

The chance of getting a blue bead is $1/p$. We expect to need $1 + 1/p$ trials including the first draw. The trick here is to let R denote the event of seeing a blue bead first, then the probability of seeing a second one is

$$1 + \mathbb{E}[R](1 + 1/p)$$

since we can treat time we draw a blue one as the starting point. R is geometrically distributed, so in total we get $1 + 1/p + 1/p^2$.

d

The probability of a group having the same number of speckles is given as $1/500^4$, so the chance of all groups having a different number of speckles is

$$(1 - 1/500^4)^{200}$$

the number we are looking for is

$$1 - (1 - 1/500^4)^{200}.$$

e

$$\frac{500}{500}\frac{499}{500}\frac{498}{500}\cdots\frac{1}{500} = \frac{500!}{500^{200}}.$$

f

There are 200 ways to put the snails in a group, and 5 orderings. The total number of ways to divide the snails is given by the binomial coefficient, so

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

in the second case we need to change to

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