

Final Examination: Solutions

1. Let \mathbf{X} and \mathbf{Y} be unit normal random variables, that is, with mean 0 and variance 1. What is the joint probability density function of $\mathbf{R} = \sqrt{\mathbf{X}^2 + \mathbf{Y}^2}$ and $\Theta = \tan^{-1}(\mathbf{Y}/\mathbf{X})$?

Solution: First of all, we have to express the change of variables and its inverse for the possible values:

$$\begin{aligned} r &= r(x, y) = \sqrt{x^2 + y^2}, \\ \theta &= \theta(x, y) = \tan^{-1}(y/x), \end{aligned}$$

and so, the inverse is given by:

$$\begin{aligned} x &= x(r, \theta) = r \cos(\theta), \\ y &= y(r, \theta) = r \sin(\theta). \end{aligned}$$

We have that the probability distribution functions of \mathbf{X} and \mathbf{Y} are given by:

$$\begin{aligned} f_{\mathbf{X}}(x) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \\ f_{\mathbf{Y}}(y) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}, \end{aligned}$$

and so, by independence of \mathbf{X} and \mathbf{Y} , we have:

$$f_{\mathbf{X}, \mathbf{Y}}(x, y) = \frac{1}{2\pi} e^{-(x^2 + y^2)/2}.$$

The formula for change of random variables in two dimensions gives us:

$$f_{\mathbf{R}, \Theta}(r, \theta) = f_{\mathbf{X}, \mathbf{Y}}(r \cos(\theta), r \sin(\theta)) \cdot \frac{1}{|\det(\frac{\partial(r, \theta)}{\partial(x, y)})|},$$

where

$$\det\left(\frac{\partial(r, \theta)}{\partial(x, y)}\right) = \det\left(\begin{array}{cc} \frac{\partial r}{\partial x} & \frac{\partial r}{\partial y} \\ \frac{\partial \theta}{\partial x} & \frac{\partial \theta}{\partial y} \end{array}\right).$$

In our case,

$$\det\left(\begin{array}{cc} \frac{\partial r}{\partial x} & \frac{\partial r}{\partial y} \\ \frac{\partial \theta}{\partial x} & \frac{\partial \theta}{\partial y} \end{array}\right) = \det\left(\begin{array}{cc} \frac{x}{r} & \frac{y}{r} \\ \frac{-y}{r^2} & \frac{x}{r^2} \end{array}\right) = \frac{x^2 + y^2}{r^3} = \frac{1}{r}.$$

Here we have used the fact you were given:

$$\frac{d}{dt} \tan^{-1}(t) = \frac{1}{1 + t^2}.$$

Putting this all together gives

$$f_{\mathbf{R}, \Theta}(r, \theta) = \frac{1}{2\pi} \cdot r \cdot e^{-r^2/2}.$$

2. Ike and Mike are shown two identical looking envelopes which have been filled with some money. One envelope has twice as much money as the other. Mike chooses an envelope at random, Ike gets the other. They are then offered the chance to change envelopes. Ike figures that with a 50% chance the other envelope has twice as much money, and 50% chance half as much, his expectation after switching should be higher. Mike thinks the same way. They switch, and each expects to do better. What is wrong with their reasoning?

Solution: Let us denote the amounts of money in the envelopes by M and $2M$. Mike's expected cash on hand after he has chosen one envelope at random of the two is

$$E(\text{Mike's envelope}) = M \cdot \frac{1}{2} + 2M \cdot \frac{1}{2} = \frac{3}{2} \cdot M.$$

Let us compute his expectation *after switching*, conditioning on the events A , that he got the envelope with the $2M$ dollars inside it, and $B = A^c$, the event that he got the envelope with only M dollars. We get:

$$\begin{aligned} E(\text{Mike's switched envelope}) &= E(\text{Mike's switched envelope}|A) \cdot P(A) \\ &\quad + E(\text{Mike's switched envelope}|A^c) \cdot P(A^c) \\ &= M \cdot \frac{1}{2} + 2M \cdot \frac{1}{2} = \frac{3}{2} \cdot M, \end{aligned}$$

that is, *the same as it was before he switched!* The error came in that although it is true that if Mike switches, he will either double his money or halve it, these two results do not start from the same case: that is, it is not possible for him to have S dollars and then be able to switch to either $\frac{1}{2}S$ or $2S$ in one move. If that *were* the case, it would make sense for Mike to switch.

[30]

3. Suppose we have N random variables $\mathbf{X} = \mathbf{X}_1, \dots, \mathbf{X}_N$, which are *i.i.d.*

(a) Define the *sample mean* $\bar{\mathbf{X}}$ of $\mathbf{X}_1, \dots, \mathbf{X}_N$. What does it mean to say this is an unbiased estimator for the mean of \mathbf{X} ?

Solution: The sample mean $\bar{\mathbf{X}}$ is defined as

$$\bar{\mathbf{X}} = \frac{1}{N} \sum_{i=1}^{i=N} \mathbf{X}_i.$$

It is an *unbiased estimator* of $\mu = E(\mathbf{X})$, the mean of \mathbf{X} , meaning that the mean $E(\bar{\mathbf{X}}) = \mu$ also. (The variance of $\bar{\mathbf{X}}$, however, is given by

$$\text{Var}(\bar{\mathbf{X}}) = \frac{1}{N} \text{Var}(\mathbf{X}),$$

so it is a much more accurate estimate of μ than \mathbf{X} itself.)

(b) The *sample variance* for $\mathbf{X}_1, \dots, \mathbf{X}_N$ is given by

$$V = \frac{1}{N-1} \sum_{i=1}^{i=N} (\mathbf{X}_i - \bar{\mathbf{X}})^2.$$

Why do you divide by $N - 1$ here and not N ? Relate this to the concept of an unbiased estimator. [No detailed calculations are required.]

Solution: As defined, \mathbf{V} has mean equal to $Var(\mathbf{X})$. (This is a calculation given in the book.) Thus it is an unbiased estimator. If we were to take the naive average of the $(\mathbf{X}_i - \bar{\mathbf{X}})^2$'s, that is, estimated $Var(\mathbf{X})$ by $\frac{1}{N} \sum (\mathbf{X}_i - \bar{\mathbf{X}})^2$, we would have made a systematic error by a factor of $\frac{N-1}{N} \neq 1$.

[25]

4. The table below gives the joint distribution of \mathbf{X} and \mathbf{Y} , which are two Bernoulli random variables.

$\frac{5}{18}$	p
$\frac{2}{9}$	$\frac{5}{18}$

(a) What is p ?

Solution: You simply have to guarantee that the sum of the terms in the grid is 1 (“law of total probability”): this gives $p = \frac{2}{9}$.

(b) What are the probability mass functions for \mathbf{X} and \mathbf{Y} ?

Solution: This just means find the two marginal distributions of the joint distribution depicted:

$\frac{1}{2}$	$\frac{5}{18}$	$\frac{2}{9}$
$\frac{1}{2}$	$\frac{2}{9}$	$\frac{5}{18}$
Y / X	$\frac{1}{2}$	$\frac{1}{2}$

That is, $P(\mathbf{X} = 1) = \frac{1}{2} = P(\mathbf{Y} = 1)$ and $P(\mathbf{X} = 0) = \frac{1}{2} = P(\mathbf{Y} = 0)$.

(c) Are \mathbf{X} and \mathbf{Y} independent?

Solution: No, because if they were independent, then the table would be given by:

$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$
$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$
Y / X	$\frac{1}{2}$	$\frac{1}{2}$

but it is not! For example,

$$P(\mathbf{X} = 0, \mathbf{Y} = 0) = \frac{2}{9} \neq \frac{1}{4} = P(\mathbf{X} = 0) \cdot P(\mathbf{Y} = 0).$$

[25]

5. One has 100 light bulbs whose lifetimes are independent exponential random variables with mean 5 hours. If the bulbs are used one at a time, with each failed bulb being replaced immediately by a new one, what is the approximate probability that there is still a working bulb after 525 hours?

Solution: This is a Central Limit Theorem problem. We let \mathbf{X}_i denote the random variable, “the lifetime of the i -th bulb”, $i = 1, \dots, 100$. The \mathbf{X}_i are *i.i.d.*'s, all having an

exponential distribution with mean 5 hours. A glance at the tables tells us that then $Var(\mathbf{X}_i) = 25$. Define

$$S_{100} = \sum_{i=1}^{i=100} \mathbf{X}_i,$$

and observe that the question can be rephrased as calculating the probability $P(S_{100} > 525)$. We can calculate this approximately using the CLT.

The CLT will compare $(S_{100} - 100 \cdot 5)/\sqrt{100 \cdot 25}$ directly with \mathbf{Z} , a standard unit normal random variable. So, we rewrite our desired condition in terms comparable to \mathbf{Z} as follows:

$$S_{100} > 525$$

is equivalent to

$$S_{100} - 500 > 25$$

which is equivalent to

$$\frac{S_{100} - 500}{\sqrt{2500}} > \frac{25}{\sqrt{2500}} = \frac{1}{2}.$$

Thus,

$$P(S_{100} > 525) \approx P(\mathbf{Z} > 0.5) = 1 - \Phi(0.5) \approx 0.3085.$$

[30]

6. Let \mathbf{X} and \mathbf{Y} be independent geometric random variables with mean T .

(a) Interpret \mathbf{X} in terms of an experiment of repeated independent coin flips? What is the bias of the coin?

Solution: A geometric random variable has possible values $\mathbf{X} = 1, \dots, n, \dots$ any positive whole number, and represents the number of that flip, in a succession of identical and independent flips, when the first head appears. If the probability that this coin will come up heads is p , then the mean is $1/p$. Thus for our case, if we know the mean is T , then we know the coin has bias, or probability of coming up heads, equal to $1/T$. Similarly for \mathbf{Y} .

(b) Interpret the random variable $\mathbf{X} + \mathbf{Y}$, and give its distribution.

Solution: Continuing the discussion from part (a), this would represent the number of the flip when the *second* head appeared. By discussion in class, we know this is given by the negative binomial distribution with parameters $r = 2, p = 1/T$, where I have followed the notation for the negative binomial distribution in the tables you were provided. More explicitly,

$$P(\mathbf{X} + \mathbf{Y} = n) = \binom{n-1}{r-1} p^r (1-p)^{n-r} = \binom{n-1}{2-1} (1/T)^2 (1 - (1/T))^{n-2}, n = 2, \dots$$

(c) What is the conditional probability $P(\mathbf{X} = i | \mathbf{X} + \mathbf{Y} = n)$? Interpret your result in light of the model in part (a).

Solution: First note that

$$P(\mathbf{X} = i | \mathbf{X} + \mathbf{Y} = n) = 0, \text{ for all } i \geq n.$$

For $i = 1, \dots, n - 1$, we have

$$\begin{aligned}
 P(\mathbf{X} = i \mid \mathbf{X} + \mathbf{Y} = n) &= \frac{P(\mathbf{X}=i, \mathbf{X}+\mathbf{Y}=n)}{P(\mathbf{X}+\mathbf{Y}=n)} \\
 &= \frac{P(\mathbf{X}=i, \mathbf{Y}=n-i)}{P(\mathbf{X}+\mathbf{Y}=n)} \\
 &= \frac{P(\mathbf{X}=i) \cdot P(\mathbf{Y}=n-i)}{P(\mathbf{X}+\mathbf{Y}=n)} \\
 &= \frac{p(1-p)^{i-1} \cdot p(1-p)^{n-i-1}}{(n-1)p^2(1-p)^{n-2}} \\
 &= \frac{1}{n-1}.
 \end{aligned}$$

Thus, the conditional probability is uniformly distributed. In other words, *once you know when the second head will be flipped*, the arrival of the first head is equally likely at any of the preceding flips.

[35]

7. If eight indistinguishable blackboards are to be divided among four schools (which *are* distinguishable), how many different ways are there for you to distribute the blackboards to the schools? If in addition you are subject to the constraint that each school must receive at least one blackboard, how many different ways are there for you to distribute the blackboards to the schools?

Solution: This uses a technique from chapter 1 (the problem is problem 31 of chapter 1 in Ross's book). Recall from chapter one that the general formula is, if you want to partition n indistinguishable objects into r *distinguishable* classes, there are

$$\binom{n+r-1}{r-1}$$

distinct ways in which to do it. See section 1.6 for more details.

Therefore, the answer to the first part of the question is just

$$\binom{8+3}{3} = \binom{11}{3} = \frac{11 \cdot 10 \cdot 9}{3 \cdot 2 \cdot 1} = 165.$$

For the second part, since four of the blackboards have already been distributed (one to each school), we only get to distribute four blackboards to the four schools, and the answer is that there are

$$\binom{4+3}{3} = \binom{7}{3} = 35$$

ways to do this.

[25]