Contents

1 hw1

1. Let $\{A_i, i \in I\}$ be a collection of sets. Prove De Morgan's Laws:

$$(\cup_i A_i)^c = \cap_i A_i^c, \ (\cap_i A_i)^c = \cup_i A_i^c.$$

Hint. The first one: $x \in (\cup_i A_i)^c \iff x \notin \cup_i A_i$ (it means x does not belong to any of them) $\iff x \notin A_i$ for any $i \iff x \in A_i^c$ for any $i \iff x \in \cap_i A_i^c$.

Answer. By definition,

$$\begin{aligned} x &\in (\cup_i A_i)^c \iff x \notin \cup_i A_i \iff x \notin A_i \text{ for any } i \in I \\ &\iff x \in A_i^c \text{ for all } i \in I \iff x \in \cap_i A_i^c, \end{aligned}$$

the first law is proved.

By the first law,

$$\cap_i A_i = \cap_i \left(A_i^c\right)^c = \left(\cup_i A_i^c\right)^c.$$

Hence

$$(\cap_i A_i)^c = \left(\left(\cup_i A_i^c\right)^c\right)^c = \cup_i A_i^c.$$

2. *a)* Let $\mathcal{F}_1$ and $\mathcal{F}_2$ be σ-fields of subsets of Ω. Show that $\mathcal{F}_1 \cap \mathcal{F}_2$, the collection of subsets of Ω that belong to both $\mathcal{F}_1$ and $\mathcal{F}_2$, is σ-field.

Hint. Verify the definition of a σ-field. Since $\Omega \in \mathcal{F}_1$ and $\Omega \in \mathcal{F}_2$, then $\Omega \in \mathcal{F}_1 \cap \mathcal{F}_2$ etc...

Answer. We check the definition of the σ-field. First $\emptyset, \Omega \in \mathcal{F}_1$ and $\emptyset, \Omega \in \mathcal{F}_2$. Hence $\emptyset, \Omega \in \mathcal{F}_1 \cap \mathcal{F}_2$.

If $A_1, A_2, \ldots \in \mathcal{F}_1 \cap \mathcal{F}_2$, then $A_1, A_2, \ldots \in \mathcal{F}_1$ and $A_1, A_2, \ldots \in \mathcal{F}_2$. Hence $\cup_n A_n \in \mathcal{F}_1$ and $\cup_n A_n \in \mathcal{F}_2$, i.e. $\cup_n A_n \in \mathcal{F}_1 \cap \mathcal{F}_2$.

If $A \in \mathcal{F}_1 \cap \mathcal{F}_2$, then $A \in \mathcal{F}_1$ and $A \in \mathcal{F}_2$. Hence $A^c \in \mathcal{F}_1$ and $A^c \in \mathcal{F}_2$, i.e. $A^c \in \mathcal{F}_1 \cap \mathcal{F}_2$

b) Let $\mathcal{A}$ be a collection of subsets of Ω, and let $\mathcal{F}_i, i \in I$, be all σ-fields that contain $\mathcal{A}$. Show that $\mathcal{F} = \cap_i \mathcal{F}_i$ is a σ-field.

Comment. b) is of interest if $\mathcal{A}$ itself is not a σ-field. Note that $\mathcal{A} \subseteq \cap_i \mathcal{F}_i$, and $\mathcal{P}(\Omega)$, the σ-field of all subsets of Ω, is among $\mathcal{F}_i$. The collection $\mathcal{F} = \cap_i \mathcal{F}_i$ is the smallest σ-field containing $\mathcal{A}$ (it is called the σ-field generated by $\mathcal{A}$).

Answer. Again, we check the definition of the σ-field; First $\emptyset, \Omega \in \mathcal{F}_i$ for all i. Hence $\emptyset, \Omega \in \cap \mathcal{F}_i$.

If $A_1, A_2, \ldots \in \cap_i \mathcal{F}_i$, then $A_1, A_2, \ldots \in \mathcal{F}_i$ for all i. Hence $\cup_n A_n \in \mathcal{F}_i$ for all i, i.e. $\cup_n A_n \in \cap \mathcal{F}_i$..

If $A \in \cap_i \mathcal{F}_i$, then $A \in \mathcal{F}_i$ for all i. Hence $A^c \in \mathcal{F}_i$ for all i, i.e. $A^c \in \cap_i \mathcal{F}_i$.

3. Suppose that eight distinct envelopes are placed at random in three distinct mailboxes.

a) In how many different ways this can be done?

Answer. The number of ways of putting eight envelopes one by one randomly into three mailboxes can be regarded as the total number of the outcomes of eight experiments with each having three possible outcomes. Hence the number of different ways is 3^8.

b) What is the probability that every mailbox receives at least one envelope? *Hint. Consider finding probability of the complementary event: "at least one mailbox is empty" = "exactly one empty" or "exactly two empty".*

Answer. Consider the event A="every mailbox receives at least one envelope". Then $\mathbf{P}(A) = 1 - \mathbf{P}(A^c)$, and

$$\mathbf{P}\left(A^c\right) = \frac{\#\left(A^c\right)}{\#\Omega}.$$

Now A^c =*at least one mailbox is empty"* = $B \cup C$ with B=*"exactly one empty"*,C = *"exactly two empty"*, and

$$\#\left(A^c\right) = (\#B) + (\#C).$$

Now, there are $\binom{3}{1} = 3$ ways to pick an empty box, and afterwards we go with 8 experiments having 2 different outcomes, there are 2^8 of them, with 2 subracted which discards the possibility of an empty box. By basic principle,

$$\#B = 3 \cdot (2^8 - 2)$$

There are $\binom{3}{2} = 3$ ways to pick two empty boxes, and afterwards all 8 envelopes go the third box:

$$\#C = 3 \cdot 1 = 3.$$

Hence

$$\begin{aligned}
\mathbf{P}\left(A^c\right) &= \frac{3 \cdot (2^8 - 2) + 3}{3^8} = \frac{85}{729}, \\
\mathbf{P}(A) &= 1 - \frac{85}{729} = \frac{644}{729} = 0.88.
\end{aligned}$$

`2nd answer (inclusion exclusion).` Let B_i ="ith box is empty". By inclusion-exclusion principle,

$$\begin{aligned}
\mathbf{P}\left(A^c\right) &= \mathbf{P}(B_1 \cup B_2 \cup B_3) = \mathbf{P}(B_1) + \mathbf{P}(B_2) + \mathbf{P}(B_3) \\
&\quad - \mathbf{P}(B_1 \cap B_2) - \mathbf{P}(B_1 \cap B_3) - \mathbf{P}(B_2 \cap B_3) \\
&= 3 \cdot \frac{2^8}{3^8} - 3 \cdot \frac{1}{3^8} = \frac{85}{729}, \\
\mathbf{P}(A) &= 1 - \frac{85}{729} = \frac{644}{729}.
\end{aligned}$$

4. How many different letter arrangements of length 4 (four letter "words") can be made using the letters MOTTO?

Answer. There are $\binom{4}{2} = 6$ distinct words with 2 O and 2 T.

There are $\binom{4}{2,1,1} = \frac{4!}{2!1!1!} = 12$ distinct words with M, 2 O's and one T, and the same number, 12, distinct words with M, 2 T's and one).

Thus we have $6 + 12 + 12 = 30$ different 4 letter words.

5. (a) In how many ways can 3 boys and 3 girls sit in a row?

Answer. There are $6! = 720$ different arrangements of 6 people in a row.

(b) In how many ways can 3 boys and 3 girls sit in a row if the boys and the girls are each to sit together?

Answer. There are 2! different arrangements of two blocks and 3! different arrangements inside of each block: $2!3!3! = 72$.

(c) In how many ways if only the boys must sit together?

Answer. There are 4! different arrangements of the boys block and 3 girls, and 3! different arrangements inside of the block: $4!3! = 144$.

(d) In how many ways if no two people of the same sex are allowed to sit together?

Answer. There are 3! arrangements of 3 girls. The boys must take two "spaces" between them (there are $3 \cdot 2 = 6$ different ways to sit two boys there), and the last boy has 2 choices (the front or the end of the row): $3!3 \cdot 2 \cdot 2 = 72$.

2 hw2

1. A woman has n keys, of which two will open her door. (a) If she tries the keys at random, discarding those that do not work,what is the probability that she will open the door on her kth try? (b) If she does not discard previously tried keys, what is the probability of no right key in k tries $(k \geq 1)$?

If $n = 9$, how many tries are needed to be 90% sure that the door is opened?

Answer. (a) There are $n\,(n-1)\ldots(n-k+1) = \frac{n!}{(n-k)!}$ equally likely orderings of k keys out of n. There are

$$\begin{aligned} & (n-2)\,(n-3)\ldots(n-2-(k-1)+1) \\ = \; & \frac{(n-2)!}{(n-2-(k-1))!} = \frac{(n-2)!}{(n-k-1)!} \end{aligned}$$

different orderings of $n-2$ wrong keys in the first $k-1$ tries, and there are 2 choices of a right key. Thus

$$\begin{aligned} \mathbf{P}\,(k\text{th try opens}) &= \frac{2 \cdot \frac{(n-2)!}{(n-k-1)!}}{\frac{n!}{(n-k)!}} = \frac{2}{n} \cdot \frac{(n-k)}{(n-1)} \\ &= \frac{2}{n} \cdot \left(1 - \frac{k-1}{n-1}\right). \end{aligned}$$

Comment. Similarly,

$$\begin{aligned} \mathbf{P}\,(\text{door closed after } k \text{ tries}) &= \frac{\frac{(n-2)!}{(n-2-k)!}}{\frac{n!}{(n-k)!}} = \frac{(n-k)\,(n-k-1)}{n\,(n-1)} \\ &= \left(1 - \frac{k}{n}\right)\left(1 - \frac{k}{n-1}\right), k = 1, \ldots, n-2. \end{aligned}$$

If $n = 9$, she can be 90% sure that the door will be opened in 6 tries.

(b) The sample space Ω is the set of all distinct results of k experiments with n outcomes: $\#\Omega = n^k$. Now,

$$\begin{aligned} & \#\,(\text{the first } k-1 \text{ tries fail}, kth \text{ succeeds}) \\ = \; & (n-2)^{k-1} \cdot 2. \end{aligned}$$

So,

$$\begin{aligned} \mathbf{P}\,(k\text{th try opens}) &= \frac{(n-2)^{k-1} \cdot 2}{n^k} = \frac{(n-2)^{k-1}}{n^{k-1}} \frac{2}{n} \\ &= \left(1 - \frac{2}{n}\right)^{k-1} \frac{2}{n}, \end{aligned}$$

and

$$\mathbf{P}\,(\text{door closed after } k \text{ tries}) = \left(1 - \frac{2}{n}\right)^k,$$

$\mathbf{P}(\text{door opened in k tries}) = 1 - \left(1 - \frac{2}{n}\right)^k$. With $n = 9$,

$$1 - \left(\frac{7}{9}\right)^k = 0.9, \left(\frac{7}{9}\right)^k = 0.1, k = \frac{\ln 0.1}{\ln \frac{7}{9}} = 10.$$

She can be 90% sure that the door will be opened in 10 tries.

Comment. All computations can be done by assuming independence (without counting) in the part (b).

2. If 4 married couples are arranged in a row, find the probability that no husband sits next to his wife. *Hint.* Let $A_k =$ "kth couple sits together", $k = 1, 2, 3, 4$. Apply inclusion/exclusion principle and count thinking about the book arrangement on the shelf:

$$\begin{aligned} & \mathbf{P}\,(\text{at least one couple sits together}) \\ = \;& \mathbf{P}\,(A_1 \cup A_2 \cup A_3 \cup A_4) = \sum_{i=1}^{4} \mathbf{P}\,(A_i) - \sum_{i<j} \mathbf{P}\left(A_i A_j\right) + \sum_{i<j<k} \mathbf{P}\left(A_i A_j A_k\right) - \mathbf{P}\,(A_1 A_2 A_3 A_4) \end{aligned}$$

For instance with $i < j$, $A_i A_j =$ "the ith couple and the jth couple go together, the remaining 4 people are arranged in any order".

Answer. The sample space Ω is the set of all possible arrangements (orderings) of 8 people: $\#\Omega = 8!$.

Considering kth couple as a block, we have 7! orderings of that block and remaining six people. Since there are $2! = 2$ different orderings inside of that block,

$$\#A_k = 7!2!.$$

For $i < j$,considering ith and jth couples as blocks, we have 6! orderings of those two blocks and remaining 4 people. Since there are $2! = 2$ different orderings inside of a block,

$$\#\left(A_i A_j\right) = 6!2!2!.$$

For $i < j < k$,considering ith, jth and kth couples as blocks, we have 5! orderings of those 3 blocks and remaining 2 people. Since there are $2! = 2$ different orderings inside of a block,

$$\#\left(A_i A_j A_k\right) = 5!2!2!2!.$$

Similarly,

$$\#\,(A_1 A_2 A_3 A_4) = 4!\,(2!)^4\,.$$

Now there are 4 couples, there are $\binom{4}{2}$ distinct pairs of 4 couples, and there are $\binom{4}{3}$ distinct triplets of 4 couples. By inclusion/exclusion principle,

$$\begin{aligned} & \mathbf{P}\,(\text{at least one couple sits together}) \\ = \;& \frac{4 \cdot 7!2! - \binom{4}{2}6!2!2! + \binom{4}{3}5!2!2!2! - 4!\,(2!)^4}{8!} = \frac{23}{35}. \end{aligned}$$

Hence

$$\mathbf{P}(\text{no husband sits next to his wife}) = 1 - \frac{23}{35} = \frac{12}{35} \approx 0.343 \;\left(e^{-1}?\right).$$

3. *Answering a question take into consideration all the information revealed before it.*

A man has five coins, two of which are double-headed, one is double-tailed, and two are normal.

He shuts his eyes, picks a coin at random, and tosses it. What is the probability that the lower face of the coin is a head?

He opens his eyes and sees that the coin is showing heads; what is the probability that the lower face is a head?

He shuts his eyes again, and tosses the coin again. What is the probability that the lower face is a head?

He opens his eyes and sees that the coin is showing heads; what is the probability that the lower face is a head?

He discards this coin, picks another at random, and tosses it. What is the probability that it shows heads? *Hint. In the previous question, you found the conditional probability that the discarded coin is double-headed.*

Answer. Let M="randomly picked coin is doubleheaded", R ="randomly picked coin is double-tailed", N ="randomly picked coin is normal", and let H_l^1="lower face is H in the first toss". By total probability law:

$$\begin{aligned}\mathbf{P}\left(H_l^1\right) &= \mathbf{P}\left(H_l^1|M\right)\mathbf{P}(M)+\mathbf{P}\left(H_l^1|R\right)\mathbf{P}(R)+\mathbf{P}\left(H_l^1|N\right)\mathbf{P}(N)\\ &= 1\cdot\frac{2}{5}+0+\frac{1}{2}\frac{2}{5}=\frac{2}{5}+\frac{1}{5}=\frac{3}{5}.\end{aligned}$$

```
He opens his eyes and sees that the coin is showing heads; what is the probability
that the lower face is a head?
```

Answer. Let H_u^1="upper face is H in the first toss". By total probability law, like above

$$\mathbf{P}\left(H_u^1\right)=\mathbf{P}\left(H_l^1\right)=\frac{3}{5},$$

and

$$\mathbf{P}\left(H_l^1|H_u^1\right)=\frac{\mathbf{P}\left(H_l^1\cap H_u^1\right)}{\mathbf{P}\left(H_u^1\right)}=\frac{\mathbf{P}(M)}{\mathbf{P}\left(H_u^1\right)}=\frac{2/5}{3/5}=2/3.$$

Comment. Note $\mathbf{P}\left(H_l^1|H_u^1\right)=\mathbf{P}\left(M|H_u^1\right)=2/3$, and $\mathbf{P}(M)=2/5$: probability of M increased.

```
He shuts his eyes again, and tosses the coin again.  What is the probability
that the lower face is a head?
```

Answer. Let H_l^2="lower face is H in the 2nd toss". Then, taking into consideration all available information,

$$\mathbf{P}\left(H_l^2|H_u^1\right)=\frac{\mathbf{P}\left(H_l^2\cap H_u^1\right)}{\mathbf{P}\left(H_u^1\right)}.$$

By total probability law,

$$\begin{aligned}\mathbf{P}\left(H_l^2\cap H_u^1\right) &= \mathbf{P}\left(H_l^2\cap H_u^1|M\right)\mathbf{P}(M)+\mathbf{P}\left(H_l^2\cap H_u^1|N\right)\mathbf{P}(N)\\ &= 1\cdot\frac{2}{5}+\frac{1}{4}\frac{2}{5}=\frac{1}{2}.\end{aligned}$$

Hence

$$\mathbf{P}\left(H_l^2|H_u^1\right) = \frac{1/2}{3/5} = 5/6.$$

`He opens his eyes and sees that the coin is showing heads; what is the probability that the lower face is a head?`

Answer. Let H_u^2="upper face is H in the 2nd toss". Then

$$\mathbf{P}\left(H_l^2|H_u^2 \cap H_u^1\right) = \frac{\mathbf{P}\left(H_l^2 \cap H_u^2 \cap H_u^1\right)}{\mathbf{P}\left(H_u^2 \cap H_u^1\right)} = \frac{\mathbf{P}(M)}{\mathbf{P}\left(H_u^2 \cap H_u^1\right)}.$$

By total probability,

$$\begin{aligned}\mathbf{P}\left(H_u^2 \cap H_u^1\right) &= \mathbf{P}\left(H_u^2 \cap H_u^1|M\right)\mathbf{P}(M) + \mathbf{P}\left(H_u^2 \cap H_u^1|N\right)\mathbf{P}(N) \\ &= \frac{2}{5} + \frac{1}{4}\frac{2}{5} = \frac{1}{2},\end{aligned}$$

and

$$\mathbf{P}\left(H_l^2|H_u^2 \cap H_u^1\right) = \frac{2/5}{1/2} = \frac{4}{5}.$$

Comment. Note $\mathbf{P}\left(H_l^2|H_u^2 \cap H_u^1\right) = \mathbf{P}\left(M|H_u^2 \cap H_u^1\right) = 4/5$, recall $\mathbf{P}\left(H_l^1|H_u^1\right) = \mathbf{P}\left(M|H_u^1\right) = 2/3$, and $\mathbf{P}(M) = 2/5$: probability of M increased again.

`He discards this coin, picks another at random, and tosses it. What is the probability that it shows heads?`

Answer. Let H_u^3="H in the toss of the second coin". Note M= "1st picked coin is doubleheaded" = "discarded coin is doubleheaded", N= 1st picked coin is normal" = "discarded coin is normal". Since we are given $H_u^1 \cap H_u^2$,

$$\mathbf{P}\left(M|H_u^2 \cap H_u^1\right) = \mathbf{P}\left(H_l^2|H_u^2 \cap H_u^1\right) = \frac{4}{5}, \mathbf{P}\left(N|H_u^2 \cap H_u^1\right) = 1 - \frac{4}{5} = \frac{1}{5},$$

and by total probability three times,

$$\begin{aligned}\mathbf{P}\left(H_u^3|H_u^2 H_u^1\right) &= \mathbf{P}\left(H_u^3|M\right)\mathbf{P}\left(M|H_u^2 H_u^1\right) + \mathbf{P}\left(H_u^3|N\right)\mathbf{P}\left(N|H_u^2 H_u^1\right) \\ &= \left(\frac{1}{4} + \frac{1}{2}\frac{2}{4}\right)\frac{4}{5} + \left(\frac{2}{4} + \frac{1}{4}\frac{1}{2}\right)\frac{1}{5} = \frac{21}{40}.\end{aligned}$$

The other way, using conditional independence,

$$\begin{aligned}\mathbf{P}\left(H_u^3|H_u^2 \cap H_u^1\right) &= \frac{\mathbf{P}\left(H_u^3 H_u^2 H_u^1\right)}{\mathbf{P}\left(H_u^2 H_u^1\right)} = \frac{\mathbf{P}\left(H_u^3 H_u^2 H_u^1|N\right)\mathbf{P}(N) + \mathbf{P}\left(H_u^3 H_u^2 H_u^1|M\right)\mathbf{P}(M)}{\mathbf{P}\left(H_u^2 H_u^1\right)} \\ &= \frac{\left(\frac{1}{2}\cdot\frac{1}{4} + \frac{1}{2}\cdot\frac{1}{2}\right)\left(\frac{1}{2}\right)^2\frac{2}{5} + \left(\frac{1}{2}\cdot\frac{2}{4} + 1\cdot\frac{1}{4}\right)\frac{2}{5}}{\left(\frac{1}{2}\right)^2\frac{2}{5} + 1\cdot\frac{2}{5}} = \frac{21}{40}.\end{aligned}$$

4. The hats of 10 people are mixed. Everybody picks up randomly a hat one by one. For any number n of people we found in class the probability $p_n = \mathbf{P}$ (no matches).

a) Find probability that only people $1, 2, 3$ (the first three) get the right hats (no matches for the rest). *Hint. If* $A_i =$ *"ith person gets the right hat", then by multiplication law,*

$$\begin{aligned}\mathbf{P}\,(\text{only } 1,2,3 \text{ get the right hats}) &= \mathbf{P}\,(1,2,3 \text{ get the right hats and no matches for the rest})\\ &= \mathbf{P}\,(\text{no matches for the rest}|A_1A_2A_3)\,\mathbf{P}\,(A_1A_2A_3)\,.\end{aligned}$$

Use p_n with an appropriate n in your answer.

Answer. Following the hint, by multiplication law (see lecture computation of $\mathbf{P}\,(A_1A_2A_3)$),

$$\begin{aligned}\mathbf{P}\,(\text{only } 1,2,3 \text{ get the right hats}) &= \mathbf{P}\,(1,2,3 \text{ get the right hats and no matches for the rest})\\ &= \mathbf{P}\,(\text{no matches for the rest}|A_1A_2A_3)\,\mathbf{P}\,(A_1A_2A_3) = p_7\frac{7!}{10!}.\end{aligned}$$

b) Find probability of exactly three matches.

Answer. We have

$$\begin{aligned}\mathbf{P}\,(\text{exactly three matches}) &= \sum_{i_1<i_2<i_3} \mathbf{P}\,(\text{only } i_1, i_2, i_3 \text{ get the right hats})\\ &= \sum_{i_1<i_2<i_3} p_7\frac{7!}{10!} = \binom{10}{3} p_7 \frac{7!}{10!} = \frac{10!}{3!7!} p_7 \frac{7!}{10!} = \frac{p_7}{3!}.\end{aligned}$$

We found in class, see note of 8/29, $p_7 = \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \frac{1}{5!} + \frac{1}{6!} - \frac{1}{7!} = \frac{103}{280}$,

$$\mathbf{P}\,(\text{exactly three matches}) = \frac{p_7}{3!} = \frac{103}{1680} \approx 6.131\,0 \times 10^{-2}.$$

Comment. We could guess the general answers: in the case of n people, for $0 \le k < n$,

$$\mathbf{P}\,(\text{exactly } k \text{ matches}) = \frac{p_{n-k}}{k!}.$$

Recall we noticed that $\lim_n p_n = e^{-1}$. Hence for large n,

$$\mathbf{P}\,(\text{exactly } k \text{ matches}) = \frac{p_{n-k}}{k!} \approx e^{-1}\frac{1}{k!},$$

that is for large n, the number X of matches is approximately Poisson with $\lambda = 1$.

5. Jane has 3 children, each of which is equally likely to be a boy or a girl independently of the others. Consider the events:

$$
\begin{aligned}
A &= \text{"all the children are of the same sex"}, \\
B &= \text{"there is at most one boy"}, \\
C &= \text{"the family includes a boy and a girl"}.
\end{aligned}
$$

(a) Show that A is independent of B, and that B is independent of C.
(b) Is A independent of C?
(c) Do these results hold if Jane has four children?
Answer. We find first

$$
\begin{aligned}
\mathbf{P}(A) &= \mathbf{P}(\text{all boys}) + \mathbf{P}(\text{all girls}) \\
&= \left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^3 = \frac{1}{4}.
\end{aligned}
$$

We could find directly

$$
\begin{aligned}
\mathbf{P}(B) &= \mathbf{P}(\text{no boys}) + \mathbf{P}(\text{exactly one boy}) \\
&= \left(\frac{1}{2}\right)^3 + 3\left(\frac{1}{2}\right)^3 = \frac{4}{8} = \frac{1}{2},
\end{aligned}
$$

because "exactly one boy"$= \{bgg, gbg, ggb\}$, or, alternatively, realize that the number X of boys is binomial with $n = 3$, $p = 1/2$, and

$$
\begin{aligned}
\mathbf{P}(B) &= \mathbf{P}(\text{at most one boy}) = \mathbf{P}(X = 0) + \mathbf{P}(X = 1) \\
&= \left(\frac{1}{2}\right)^3 + \binom{3}{1}\left(\frac{1}{2}\right)^3 = \frac{1}{2}.
\end{aligned}
$$

Now, $C = A^c$, and

$$
\mathbf{P}(C) = 1 - \mathbf{P}(A) = 1 - \frac{1}{4} = \frac{3}{4}.
$$

(a) We have

$$
\begin{aligned}
\mathbf{P}(A \cap B) &= \mathbf{P}(\text{all boys}) = \left(\frac{1}{2}\right)^3 = \frac{1}{8} \\
&= \frac{1}{4} \cdot \frac{1}{2} = \mathbf{P}(A)\,\mathbf{P}(B),
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbf{P}(B \cap C) &= \mathbf{P}(\text{exactly one boy}) = 3\left(\frac{1}{2}\right)^3 = \frac{3}{8} \\
&= \frac{1}{2}\frac{3}{4} = \mathbf{P}(B)\,\mathbf{P}(C).
\end{aligned}
$$

That is A is independent of B, and B is independent of C.

(b) No,

$$\mathbf{P}(A\cap C)=\mathbf{P}(\emptyset)=0\neq\mathbf{P}(A)\,\mathbf{P}(C)\,.$$

(c) No, in that case,

$$\begin{aligned}\mathbf{P}(A) &= \mathbf{P}(\text{all boys})+\mathbf{P}(\text{all girls})\\ &= \left(\frac{1}{2}\right)^4+\left(\frac{1}{2}\right)^4=\frac{1}{8},\end{aligned}$$

$$\begin{aligned}\mathbf{P}(B) &= \mathbf{P}(\text{no boys})+\mathbf{P}(\text{exactly one boy})\\ &= \left(\frac{1}{2}\right)^4+4\left(\frac{1}{2}\right)^4=\frac{5}{16},\end{aligned}$$

and

$$\mathbf{P}(C)=1-\mathbf{P}(A)=1-\frac{1}{8}=\frac{7}{8}.$$

On the other hand,

$$\begin{aligned}\mathbf{P}(A\cap B) &= \mathbf{P}(\text{all boys})=\left(\frac{1}{2}\right)^4=\frac{1}{16}\\ &\neq \mathbf{P}(A)\,\mathbf{P}(B)\,,\end{aligned}$$

and

$$\mathbf{P}(B\cap C)=\mathbf{P}(\text{exactly one boy})=4\left(\frac{1}{2}\right)^4=\frac{1}{4}\neq\mathbf{P}(B)\,\mathbf{P}(C)\,.$$

3 hw3

1. In each packet of Corn Flakes may be found a plastic bust of one of the last five Vice-Chancellors of Cambridge University, the probability that any given packet contains any specific Vice-Chancellor being 1/5, independently of all other packets. Show that the probability that each of the last three Vice-Chancellors is obtained in a bulk purchase of six packets is

$$1-3\left(\frac{4}{5}\right)^6+3\left(\frac{3}{5}\right)^6-\left(\frac{2}{5}\right)^6.$$

Hint. Let A_k="kth chancellor is in at least one of the 6 packets", $k=1,\ldots,5$. Then, by de Morgan law and inclusion-exclusion principle,

$$\begin{aligned}
\mathbf{P}\left(A_3\cap A_4\cap A_5\right) &= 1-\mathbf{P}\left(\left(A_3\cap A_4\cap A_5\right)^c\right)=1-\mathbf{P}\left(A_3^c\cup A_4^c\cup A_5^c\right),\\
\mathbf{P}\left(A_3^c\cup A_4^c\cup A_5^c\right) &= \sum_{k=3}^{5}\mathbf{P}\left(A_k^c\right)-\sum_{3\le i<j\le 5}\mathbf{P}\left(A_i^c\cap A_j^c\right)+\mathbf{P}\left(A_3^c\cap A_4^c\cap A_5^c\right).
\end{aligned}$$

Answer. Let A_k="kth chancellor is in at least one of the 6 packets", $k=1,\ldots,5$. Then, by de Morgan law and inclusion-exclusion principle,

$$\begin{aligned}
\mathbf{P}\left(A_3\cap A_4\cap A_5\right) &= 1-\mathbf{P}\left(\left(A_3\cap A_4\cap A_5\right)^c\right)=1-\mathbf{P}\left(A_3^c\cup A_4^c\cup A_5^c\right),\\
\mathbf{P}\left(A_3^c\cup A_4^c\cup A_5^c\right) &= \sum_{k=3}^{5}\mathbf{P}\left(A_k^c\right)-\sum_{3\le i<j\le 5}\mathbf{P}\left(A_i^c\cap A_j^c\right)+\mathbf{P}\left(A_3^c\cap A_4^c\cap A_5^c\right).
\end{aligned}$$

Since the probability that any given packet does not contain a specific Vice-Chancellor is 4/5, independently of all other packets,

$$\mathbf{P}\left(A_k^c\right)=\left(\frac{4}{5}\right)^6=\frac{4^6}{5^6}, k=3,4,5;$$

Since the probability that any given packet does not contain two specific Vice-Chancellors is 3/5, independently of all other packets,

$$\mathbf{P}\left(A_i^c\cap A_j^c\right)=\left(\frac{3}{5}\right)^6=\frac{3^6}{5^6}, 3\le i<j\le 5;$$

Since the probability that any given packet does not contain three specific Vice-Chancellors is 2/5, independently of all other packets,

$$\mathbf{P}\left(A_3^c\cap A_4^c\cap A_5^c\right)=\left(\frac{2}{5}\right)^6=\frac{2^6}{5^6}.$$

Hence

$$\begin{aligned}
\mathbf{P}\left(A_3^c\cup A_4^c\cup A_5^c\right) &= 3\cdot\left(\frac{4}{5}\right)^6-\binom{3}{2}\left(\frac{3}{5}\right)^6+\left(\frac{2}{5}\right)^6,\\
\mathbf{P}\left(A_3\cap A_4\cap A_5\right) &= 1-3\left(\frac{4}{5}\right)^6+3\left(\frac{3}{5}\right)^6-\left(\frac{2}{5}\right)^6=0.349\,44.
\end{aligned}$$

2. There are two roads from A to B and two roads from B to C. Each of the four roads is blocked by snow with probability p, independently of the others. Find the probability that there is an open road from A to B given that there is no open route from A to C.

If, in addition, there is a direct road from A to C, this road being blocked with probability p independently of the others, find the required conditional probability.

Answer. Denote AB="A to B open", $\overline{AB}$="A to B closed", BC="B to C open", $\overline{BC}$="B to C closed", AC="A to C open", $\overline{AC}$="A to C closed". Then

$$\mathbf{P}\left(AB|\overline{AC}\right)=\frac{\mathbf{P}\left(AB\cap\overline{AC}\right)}{\mathbf{P}\left(\overline{AC}\right)}=\frac{\mathbf{P}\left(AB\cap\overline{AC}\right)}{\mathbf{P}\left(\overline{AC}\right)}=\frac{\mathbf{P}\left(AB\right)\mathbf{P}\left(\overline{BC}\right)}{\mathbf{P}\left(\overline{AB}\cup\overline{BC}\right)}.$$

Using independence and inclusion-exclusion principle,

$$\begin{aligned}\mathbf{P}\left(AB\right) &= 1-\mathbf{P}\left(\overline{AB}\right)=1-p^2,\mathbf{P}\left(\overline{BC}\right)=p^2,\\ \mathbf{P}\left(\overline{AB}\cup\overline{BC}\right) &= \mathbf{P}\left(\overline{AB}\right)+\mathbf{P}\left(\overline{BC}\right)-\mathbf{P}\left(\overline{AB}\cap\overline{BC}\right)=p^2+p^2-p^2\cdot p^2=(2-p^2)p^2.\end{aligned}$$

Hence

$$\mathbf{P}\left(AB|\overline{AC}\right)=\frac{\left(1-p^2\right)p^2}{\left(2-p^2\right)p^2}=\frac{1-p^2}{2-p^2}.$$

Assume there is direct new road from A to C. Denote N="new road open", $\bar{N}$="new road closed". Then, using independence, and previous computation,

$$\begin{aligned}\mathbf{P}\left(AB|\overline{AC}\right) &= \frac{\mathbf{P}\left(AB\cap\overline{AC}\right)}{\mathbf{P}\left(\overline{AC}\right)}=\frac{\mathbf{P}\left(AB\cap\overline{BC}\cap\bar{N}\right)}{\mathbf{P}\left(\bar{N}\cap\left(\overline{AB}\cup\overline{BC}\right)\right)}=\frac{\mathbf{P}\left(AB\right)\mathbf{P}\left(\overline{BC}\right)\mathbf{P}\left(\bar{N}\right)}{\mathbf{P}\left(\bar{N}\right)\mathbf{P}\left(\overline{AB}\cup\overline{BC}\right)}\\ &= \frac{\mathbf{P}\left(AB\right)\mathbf{P}\left(\overline{BC}\right)}{\mathbf{P}\left(\overline{AB}\cup\overline{BC}\right)}=\frac{1-p^2}{2-p^2}.\end{aligned}$$

Alternatively (and there are other alternatives), by independence and total probability law,

$$\begin{aligned}\mathbf{P}\left(AB|\overline{AC}\right) &= \frac{\mathbf{P}\left(AB\cap\overline{AC}\right)}{\mathbf{P}\left(\overline{AC}\right)}=\frac{\mathbf{P}\left(AB\cap\overline{AC}|N\right)\mathbf{P}\left(N\right)+\mathbf{P}\left(AB\cap\overline{AC}|\bar{N}\right)\mathbf{P}\left(\bar{N}\right)}{\mathbf{P}\left(\overline{AC}|N\right)\mathbf{P}\left(N\right)+\mathbf{P}\left(\overline{AC}|\bar{N}\right)\mathbf{P}\left(\bar{N}\right)}\\ &= \frac{\mathbf{P}\left(AB\cap\overline{AC}|\bar{N}\right)\mathbf{P}\left(\bar{N}\right)}{\mathbf{P}\left(\overline{AC}|\bar{N}\right)\mathbf{P}\left(\bar{N}\right)}=\frac{\mathbf{P}\left(AB\cap\overline{AC}|\bar{N}\right)}{\mathbf{P}\left(\overline{AC}|\bar{N}\right)}=\frac{1-p^2}{2-p^2}.\end{aligned}$$

3. Consider a gambler G ruin problem where he starts with k dollars, $0<k<N$. A fair coin is tossed repeatedly. G wins \$1 if H, and loses \$1 if T. The game stops in two cases: either G is ruined or G reaches the desired amount \$$N$. Show that the game stops with probability 1. Hint. Besides the ruin probability p_k consider the probability $\bar{p}_k$ to reach N. Besides the equation (6) on p.17 for p_k, write a similar equation for $\bar{p}_k$ and solve it (what are $\bar{p}_0,\bar{p}_N$?). Find $p_k+\bar{p}_k$.

Answer. Let A="1st toss is H", B_k="G reaches N with initial k dollars". By total probability law,

$$\begin{aligned}\bar{p}_k &= \mathbf{P}\left(B_k\right)=\mathbf{P}\left(B_k|A\right)\mathbf{P}\left(A\right)+\mathbf{P}\left(B_k|A^c\right)\mathbf{P}\left(A^c\right)\\ &= \mathbf{P}\left(B_{k+1}\right)\frac{1}{2}+\mathbf{P}\left(B_{k-1}\right)\frac{1}{2}=\frac{1}{2}\bar{p}_{k+1}+\frac{1}{2}\bar{p}_k,\end{aligned}$$

that is equations for p_k and $\bar{p}_k$ are the same. In both cases,

$$\begin{aligned}\frac{1}{2}(\bar{p}_{k+1}-\bar{p}_k) &= \frac{1}{2}(\bar{p}_k-\bar{p}_{k-1}), k\geq 1,\\ \frac{1}{2}(p_{k+1}-p_k) &= \frac{1}{2}(p_k-p_{k-1}), k\geq 1,\end{aligned}$$

i.e., the differences are constant: $\bar{p}_k-\bar{p}_{k-1}=\bar{d}$, $p_k-p_{k-1}=d, k\geq 1$. Hence

$$\begin{aligned}p_k &= p_{k-1}+d=\ldots=p_0+kd,\\ \bar{p}_k &= \bar{p}_{k-1}+\bar{d}=\ldots=\bar{p}_0+k\bar{d}.\end{aligned}$$

Since $p_0=\bar{p}_N=1, p_N=\bar{p}_0=0$, we find that $p_k=1-k/N, \bar{p}_k=k/N, k\geq 0$. Let C_k="G is ruined with initial k dollars". Since C_k and B_k are mutually exclusive (disjoint),

$$\mathbf{P}(B_k\cup C_k)=\mathbf{P}(B_k)+\mathbf{P}(C_k)=\bar{p}_k+p_k=1-k/N+k/N=1,$$

that is the game stops with probability 1.

4. A communication channel transmits a signal as sequence of digits 0 and 1. The probability of incorrect reception of each digit is p. To reduce the probability of error at reception, 0 is transmitted as 00000 (five zeroes) and 1, as 11111. Assume that the digits are received independently and the majority decoding is used. Compute the probability of receiving the signal incorrectly if the original signal is (a) 0; (b) 101. Evaluate the probabilities when $p=0.2$. Hint. Number of errors in a reception of 5 digits is binomial r.v.

Answer. (a) X, the number of errors in 00000 transmission, is binomial$(5,p)$. Then, denoting $a=\mathbf{P}$ (0 is received incorrectly) ,

$$\begin{aligned}a &= \mathbf{P}\,(0 \text{ is received incorrectly})\\ &= \mathbf{P}(X\geq 3)=f(3)+f(4)+f(5)\\ &= \binom{5}{3}p^3(1-p)^2+\binom{5}{4}p^4(1-p)+\binom{5}{5}p^5\\ &= 10p^3\left(1-2p+p^2\right)+5p^4(1-p)+p^5\\ &= p^3\left(6p^2-15p+10\right).\end{aligned}$$

If $p=0.2$, then $a=0.2^3\left(6\cdot 0.2^2-15\cdot 0.2+10\right)=0.057\,92$. The same answer is obtained for $\mathbf{P}$ (1 is received incorrectly) .

Comment. The new error $a<p$ considerably.

(b) Now, "101 is received incorrectly"= "at least one of the digits in 101 is received incorrectly", and by part (a),

$$\begin{aligned}&\mathbf{P}\,(\text{digit is received correctly})\\ &= 1-a=1-p^3\left(6p^2-15p+10\right).\end{aligned}$$

By independence, with $a = p^3 \left(6p^2 - 15p + 10\right)$,

$$\begin{aligned} & \mathbf{P}\,(\text{101 is received incorrectly}) \\ = \;& 1 - \mathbf{P}\,(\text{all three received correctly}) \\ = \;& 1 - (1-a)^3\,. \end{aligned}$$

If $p = 0.2$, then

$$\begin{aligned} & \mathbf{P}\,(\text{101 is received incorrectly}) \\ = \;& 1 - (1 - 0.057\,92)^3 = 0.163\,89. \end{aligned}$$

5. A coin with $\mathbf{P}(H) = p, \mathbf{P}(T) = q = 1 - p$, is tossed repeatedly (indefinitely). Let $H_k =$ "H in the kth toss", $T_k =$ "T in the kth toss". Assume all tosses are independent.

(a) Find $\mathbf{P}\,(\text{at least one } H \text{ after } n) = \mathbf{P}\left(\cup_{m=n}^{\infty} H_m\right) = 1 - \mathbf{P}\left(\cap_{m=n}^{\infty} T_m\right)$.

Hint. Recall, by continuity of probability, $\mathbf{P}\left(\cap_{m=n}^{\infty} T_m\right) = \lim_{l\to\infty} \mathbf{P}\left(\cap_{m=n}^{n+l} T_m\right)$.

Answer. Indeed,

$$\mathbf{P}\,(\text{at least one } H \text{ after } n) = \mathbf{P}\left(\cup_{m=n}^{\infty} H_m\right) = 1 - \mathbf{P}\left(\cap_{m=n}^{\infty} T_m\right),$$

and by continuity of probability and independence,

$$\mathbf{P}\left(\cap_{m=n}^{\infty} T_m\right) = \lim_{l\to\infty} \mathbf{P}\left(\cap_{m=n}^{n+l} T_m\right) = \lim_{l\to\infty} q^{l+1} = 0.$$

Therefore

$$\mathbf{P}\,(\text{at least one } H \text{ after } n) = \mathbf{P}\left(\cup_{m=n}^{\infty} H_m\right) = 1 - \mathbf{P}\left(\cap_{m=n}^{\infty} T_m\right) = 1.$$

(b) Find probability of infinitely many H's.

Hint. $\mathbf{P}\,(\text{infinitely many } H) = \mathbf{P}\left(\cap_{n=1}^{\infty} \cup_{m=n}^{\infty} H_m\right)$: use the previous part (a) and continuity of probability.

Answer. By part (a) and continuity of probability, $B_n = \cup_{m=n}^{\infty} H_m$ is decreasing sequence of events,

$$\begin{aligned} \mathbf{P}\left(\text{infinitely many } H's\right) &= \mathbf{P}\left(\cap_{n=1}^{\infty} \cup_{m=n}^{\infty} H_m\right) \\ &= \lim_{n\to\infty} \mathbf{P}\left(\cup_{m=n}^{\infty} H_m\right) = 1. \end{aligned}$$

4 hw4

1. Consider the following strategy playing the roulette. Bet \$1 on red. If red appears (which happens with probability 18/38), then take \$1 and stop playing for the day. If red does not appear, then bet additional \$1 on red each of the following two rounds, and then stop playing for the day no matter the outcome. Let X be the net gain (a negative gain means a loss).

(a) What are possible values of X? Find $\mathbf{P}(X = k)$ for all possible values k; sketch the distribution function (cdf) of X.Hint. To find all $\mathbf{P}(X = k)$ it is convenient to use probability tree.

(b) Compute $\mathbf{P}(X > 0)$, the probability of net win. Is it a good strategy? (c) You played 5 days. Find probability to win \$1 in at least three of 5 games.

Answer. (a) The range of $X = \{1, -1, -3\}$. Using tree diagram or directly,

$$\begin{aligned}\mathbf{P}(X = 1) &= \mathbf{P}(W_1) + \mathbf{P}(L_1 W_2 W_3)\\ &= \frac{18}{38} + \left(\frac{18}{38}\right)^2 \frac{20}{38} = \frac{4059}{6859} = 0.591\,78,\end{aligned}$$

$$\begin{aligned}\mathbf{P}(X = -1) &= \mathbf{P}(L_1 W_2 L_3) + \mathbf{P}(L_1 L_2 W_3)\\ &= 2 \cdot \left(\frac{20}{38}\right)^2 \frac{18}{38} = \frac{1800}{6859} = 0.262\,43,\end{aligned}$$

and

$$\mathbf{P}(X = -3) = \mathbf{P}(L_1 L_2 L_3) = \left(\frac{20}{38}\right)^3 = \frac{1000}{6859} = 0.145\,79.$$

(b) By (a), probability of a positive net gain is

$$\mathbf{P}(X > 0) = \mathbf{P}(X = 1) = \frac{4059}{6859} = 0.591\,78.$$

Although $\mathbf{P}(X > 0) > 0.5$, but positive net gain is \$1 at most, and the net loss could reach \$3. We will see later that average net gain per day (if we play many games using this strategy) is negative ($\approx$ 11 cents).

(c) $Y = \#$ of days you won \$1 is binomial$\left(n = 5, p = \frac{4059}{6859} = 0.591\,78\right)$. Hence

$$\begin{aligned}\mathbf{P}(Y \geq 3) &= \sum_{k=3}^{5} \mathbf{P}(Y = k) = \sum_{k=3}^{5} \binom{5}{k} \left(\frac{4059}{6859}\right)^k \left(1 - \frac{4059}{6859}\right)^{5-k}\\ &= \binom{5}{3} \left(\frac{4059}{6859}\right)^3 \left(1 - \frac{4059}{6859}\right)^2 + \binom{5}{4} \left(\frac{4059}{6859}\right)^4 \left(1 - \frac{4059}{6859}\right) + \left(\frac{4059}{6859}\right)^5\\ &= 0.668\,26.\end{aligned}$$

2. a) Let U be a r.v. with distribution function

$$F_U(x) = \mathbf{P}(U \leq u) = \begin{cases} 0 & \text{if } u < 0, \\ u & \text{if } 0 \leq u \leq 1, \\ 1 & \text{if } u > 1. \end{cases}$$

We say U is uniformly distributed in the interval $[0,1]$; note that

$$\mathbf{P}(a < U \le b) = \frac{b-a}{1} = b-a, 0 \le a < b \le 1.$$

Let F be a distribution function which is continuous and strictly increasing (note that range of F is $(0,1)$). Recall that in this case the inverse function $F^{-1} : (0,1) \to \mathbf{R}$ is strictly increasing continuous and cancellation identities hold: $F\left(F^{-1}(x)\right) = x, x \in (0,1)$, and $F^{-1}(F(x)) = x, x \in \mathbf{R}$.

a) Show that $X = F^{-1}(U)$ is a r.v. having distribution function F, i.e. $F = F_X$.

Answer. Since F is continuous and strictly increasing, $\left\{F^{-1}(U) \le x\right\} = \left\{F\left(F^{-1}(U)\right) \le F(x)\right\} = \{U \le F(x)\}$, and $F(x) \in (0,1)$ for any $x \in \mathbf{R}$. Hence

$$\begin{aligned}
\mathbf{P}(X \le x) &= \mathbf{P}\left(F^{-1}(U) \le x\right) = \mathbf{P}\left(F\left(F^{-1}(U)\right) \le F(x)\right) \\
&= \mathbf{P}(U \le F(x)) = F(x), x \in \mathbf{R}.
\end{aligned}$$

b) Let X be a r.v. with a continuous distribution function F. Find expression for the distribution functions of the following random variables:

(i) X^2; (ii) $|X|$ and $\sqrt{|X|}$; (iii) $F(X)$, assuming F is strictly increasing; (iv) $G^{-1}(F(X))$, assuming F and $G : \mathbf{R} \to (0,1)$ are strictly increasing with $G^{-1} : (0,1) \to \mathbf{R}$ as the inverse of G.

Answer. (i) Since $X^2 \ge 0$, we have $\mathbf{P}\left(X^2 \le y\right) = 0$ if $y < 0$. Hence

$$\begin{aligned}
\mathbf{P}\left(X^2 \le y\right) &= \mathbf{P}\left(|X| \le \sqrt{y}\right) = \mathbf{P}\left(-\sqrt{y} \le X \le \sqrt{y}\right) \\
&= F\left(\sqrt{y}\right) - F\left(-\sqrt{y}\right), y \ge 0,
\end{aligned}$$

because F is continuous. That is

$$F_{X^2}(y) = \begin{cases} 0 & \text{if } y < 0, \\ F\left(\sqrt{y}\right) - F\left(-\sqrt{y}\right) & \text{if } y \ge 0. \end{cases}$$

(ii) For $y < 0$ that $\mathbf{P}(|X| \le y) = 0$. For $y \ge 0$, $\{|X| \le y\} = \{-y \le X \le y\}$, and

$$\mathbf{P}(|X| \le y) = \mathbf{P}(-y \le X \le y) = F(y) - F(-y).$$

So,

$$F_{|X|}(y) = \begin{cases} 0 & \text{if } y < 0, \\ F(y) - F(-y) & \text{if } y \ge 0. \end{cases}$$

For $y < 0$ that $\mathbf{P}\left(\sqrt{|X|} \le y\right) = 0$. For $y \ge 0$, $\left\{\sqrt{|X|} \le y\right\} = \left\{|X| \le y^2\right\}$, and

$$\mathbf{P}\left(\sqrt{|X|} \le y\right) = \mathbf{P}\left(|X| \le y^2\right) = F\left(y^2\right) - F\left(-y^2\right).$$

Hence

$$F_{\sqrt{|X|}}(y) = \begin{cases} 0 & \text{if } y < 0, \\ F\left(y^2\right) - F\left(-y^2\right) & \text{if } y \ge 0. \end{cases}$$

(iii) Let $Y = F(X)$. Since F is strictly increasing, its range is $(0,1)$, and $\mathbf{P}(Y \le y) = 0$ if $y < 0$, $\mathbf{P}(Y \le y) = 1$ if $y \ge 1$. For $y \in (0,1)$,

$$\begin{aligned}
\mathbf{P}(F(X) \le y) &= \mathbf{P}\left(F^{-1}(F(X)) \le F^{-1}(y)\right) \\
&= \mathbf{P}\left(X \le F^{-1}(y)\right) = F\left(F^{-1}(y)\right) = y.
\end{aligned}$$

Hence

$$F_Y(y) = \begin{cases} 0 & \text{if } y \le 0, \\ y & \text{if } 0 < y < 1, \\ 1 & \text{if } y \ge 1; \end{cases}$$

Y is a uniform r.v. in $(0, 1)$.

(iv) We have $\left\{G^{-1}(F(X)) \le y\right\} = \{F(X) \le G(y)\}$, $y \in \mathbf{R}$, and, according to (iii), $F(X)$ is uniform. So,

$$\mathbf{P}(F(X) \le G(y)) = G(y), \, y \in \mathbf{R}.$$

3. A coin is tossed repeatedly and heads turns up on each toss with probability p. Let H_n and T_n be the numbers of heads and tails in n tosses. Show that for each $\varepsilon > 0$,

$$\mathbf{P}\left(2p - 1 - \varepsilon \le \frac{1}{n}(H_n - T_n) \le 2p - 1 + \varepsilon\right) \to 1$$

as $n \to \infty$. Hint. $H_n + T_n = n$.

Answer. Since $H_n + T_n = n$, and $T_n = n - H_n$, we have

$$\begin{aligned} & \left\{2p - 1 - \varepsilon \le \frac{1}{n}(H_n - T_n) \le 2p - 1 + \varepsilon\right\} \\ = & \left\{2p - 1 - \varepsilon \le \frac{1}{n}(2H_n - n) \le 2p - 1 + \varepsilon\right\} \\ = & \left\{2p - 1 - \varepsilon \le \left(2\frac{H_n}{n} - 1\right) \le 2p - 1 + \varepsilon\right\} \\ = & \left\{2p - \varepsilon \le 2\frac{H_n}{n} \le 2p + \varepsilon\right\} = \left\{-\varepsilon \le 2\left(\frac{H_n}{n} - p\right) \le \varepsilon\right\} \\ = & \left\{-\frac{\varepsilon}{2} \le \frac{H_n}{n} - p \le \frac{\varepsilon}{2}\right\}. \end{aligned}$$

Applying Bernoulli theorem with $\varepsilon/2$ (hat theorem holds for all $\varepsilon > 0$),

$$\begin{aligned} & \mathbf{P}\left(2p - 1 - \varepsilon \le \frac{1}{n}(H_n - T_n) \le 2p - 1 + \varepsilon\right) \\ = & \mathbf{P}\left(-\frac{\varepsilon}{2} \le \frac{H_n}{n} - p \le \frac{\varepsilon}{2}\right) \to 0 \end{aligned}$$

as $n \to \infty$.

2nd answer. As stated and discussed in class, $H_n/n \to p$ as $n \to \infty$ with probability 1. Then

$$\frac{1}{n}(H_n - T_n) = \frac{1}{n}[H_n - (n - H_n)] = \frac{2H_n - n}{n} = 2 \cdot \frac{H_n}{n} - 1 \to 2p - 1$$

as $n \to \infty$ with probability 1.

3rd answer. As stated and discussed in class, $H_n/n \to p$, and $T_n/n \to q = 1 - p$ as $n \to \infty$ with probability 1. Then

$$\frac{1}{n}(H_n - T_n) = \frac{H_n}{n} - \frac{T_n}{n} \to p - (1 - p) = 2p - 1$$

as $n \to \infty$ with probability 1.

Comment. The convergence with probability 1 is stronger. If $X_n \to a$ as $n \to \infty$ with probability 1, then for each $\varepsilon > 0$, by continuity of probability,

$$\begin{aligned} 1 &= \mathbf{P}\left(\cup_{n=1}^{\infty} \cap_{k=n}^{\infty} \{-\varepsilon < X_k - a < \varepsilon\}\right) = \lim_{n\to\infty} \mathbf{P}\left(\cap_{k=n}^{\infty} \{-\varepsilon < X_k - a < \varepsilon\}\right) \\ &\leq \lim_{n\to\infty} \mathbf{P}\left(\{-\varepsilon < X_n - a < \varepsilon\}\right). \end{aligned}$$

Thus $\lim_{n\to\infty} \mathbf{P}\left(\{-\varepsilon < X_n - a < \varepsilon\}\right) = 1$.

4. (Binomial distribution as an approximation) An urn contains N balls, m of which are red. A random sample of n balls is withdrawn without replacement from the urn. Assume $n < m < N$. The number X of red balls in this sample has the mass function

$$\mathbf{P}(X = k) = \frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}}, k = 0, \ldots, n.$$

This is called the hypergeometric distribution with parameters N, m, n, and X is a number of "successes" in n trials with initial "success" probability m/N but the trials are not independent.

Show further that if N, m approach ∞ in such a way that $m/N \approx p \in (0,1)$, then for $k = 0, \ldots, n$,

$$\mathbf{P}(X = k) \approx \binom{n}{k} p^k (1-p)^{n-k}$$

Comment. It shows that for fixed n and large N, m, the distribution of X is approximately binomial$\left(n, p = \frac{m}{N}\right)$, the trials for large m, N with $m/N \approx p$ are approximately independent: the distribution in that case barely depends on whether or not the balls are replaced in the urn immediately after their withdrawal.

Hint. Write explicitly the binomials, divide numerator and denominator by N^n and group:

$$\begin{aligned} &\mathbf{P}(X = k) \\ = \ &\frac{n!}{k!\,(n-k)!} \frac{\frac{m(m-1)\ldots(m-k+1)}{N^k} \frac{(N-m)(N-m-1)\ldots(N-m-(n-k)+1)}{N^{n-k}}}{\frac{N(N-1)\ldots(N-n+1)}{N^n}} \end{aligned}$$

Answer. Let $k \leq n \leq m \leq N$. Then we group the following way:

$$\begin{aligned} &\mathbf{P}(X = k) \\ = \ &\frac{n!}{k!\,(n-k)!} \frac{\frac{m(m-1)\ldots(m-k+1)}{N^k} \frac{(N-m)(N-m-1)\ldots(N-m-(n-k)+1)}{N^{n-k}}}{\frac{N(N-1)\ldots(N-n+1)}{N^n}} \end{aligned}$$

As $N, m \to \infty$, and $m/N \to p \in (0,1)$, and $k \le n$ remain fixed, we have

$$\begin{aligned}
\mathbf{P}(X=k) &= \frac{n!}{k!\,(n-k)!}\frac{\frac{m}{N}\left(\frac{m}{N}-\frac{1}{N}\right)\dots\left(\frac{m}{N}-\frac{k-1}{N}\right)\left(1-\frac{m}{N}\right)\left(1-\frac{m}{N}-\frac{1}{N}\right)\dots\left(1-\frac{m}{N}-\frac{n-k-1}{N}\right)}{\left(1-\frac{1}{N}\right)\dots\left(1-\frac{n-1}{N}\right)} \\
&\approx \frac{n!}{k!\,(n-k)!}\frac{m}{N}\left(\frac{m}{N}-\frac{1}{N}\right)\dots\left(\frac{m}{N}-\frac{k-1}{N}\right)\times \\
&\quad\times\left(1-\frac{m}{N}\right)\left(1-\frac{m}{N}-\frac{1}{N}\right)\dots\left(1-\frac{m}{N}-\frac{n-k-1}{N}\right) \\
&\approx \binom{n}{k}\left(\frac{m}{N}\right)^k\left(1-\frac{m}{N}\right)^{n-k}
\end{aligned}$$

Hence approximately, X is binomial $\left(n, p=\frac{m}{N}\right)$.

5. Let X be binomial$(n, p=1/2)$. It can be proved that the most likely values of X are

$$\frac{n}{2} \text{ if } n \text{ is even,}$$
$$\frac{n-1}{2} \text{ and } \frac{n+1}{2} \text{ if } n \text{ is odd.}$$

Show that *for large* n,

$$\begin{aligned}
\mathbf{P}\left(X=\frac{n}{2}\right) &= \binom{n}{n/2}2^{-n} \approx \frac{1}{\sqrt{\pi\cdot\frac{n}{2}}} \text{ if } n \text{ is even,} \\
\mathbf{P}\left(X=\frac{n+1}{2}\right) &= \binom{n}{(n+1)/2}2^{-n} \approx \frac{1}{\sqrt{\pi\cdot\frac{n}{2}}} \text{ if } n \text{ is odd.}
\end{aligned}$$

Hint. Use Stirling formula according to which

$$k! \approx \sqrt{2\pi k}\left(\frac{k}{e}\right)^k \text{ for large } k,$$

more precisely

$$\lim_{k\to\infty}\frac{k!}{\sqrt{2\pi k}\left(\frac{k}{e}\right)^k} = 1.$$

Answer. For an odd n, using Stirling formula,

$$\begin{aligned}
& \mathbf{P}\left(X=\frac{n+1}{2}\right) \\
= \quad & \binom{n}{(n+1)/2} 2^{-n}=\frac{n!}{\left(\frac{n+1}{2}\right)!\left(\frac{n-1}{2}\right)!} 2^{-n} \\
\approx \quad & \frac{\sqrt{2\pi n}\left(\frac{n}{e}\right)^{n}}{\sqrt{2\pi\frac{n+1}{2}}\left(\frac{n+1}{2e}\right)^{\frac{n+1}{2}} \sqrt{2\pi\frac{n-1}{2}}\left(\frac{n-1}{2e}\right)^{\frac{n-1}{2}}} 2^{-n} \\
= \quad & \frac{\sqrt{2\pi n} n^{n}}{\sqrt{\pi(n+1)}(n+1)^{\frac{n+1}{2}} \sqrt{\pi(n-1)}(n-1)^{\frac{n-1}{2}}} \\
= \quad & \frac{\sqrt{2\pi n}}{\sqrt{\pi(n+1)}\left(1+\frac{1}{n}\right)^{\frac{n+1}{2}} \sqrt{\pi(n-1)}\left(1-\frac{1}{n}\right)^{\frac{n-1}{2}}} \\
= \quad & \frac{1}{\sqrt{\frac{\pi n}{2}\left(1+\frac{1}{n}\right)}\left(1+\frac{1}{n}\right)^{\frac{n+1}{2}} \sqrt{1-\frac{1}{n}}\left(1-\frac{1}{n}\right)^{\frac{n-1}{2}}} \approx \frac{1}{\sqrt{\frac{\pi n}{2}}\left(1+\frac{1}{n}\right)^{\frac{n}{2}}\left(1-\frac{1}{n}\right)^{\frac{n}{2}}} \\
\approx \quad & \frac{1}{\sqrt{\frac{\pi n}{2}} e^{1/2} e^{-1/2}}=\frac{1}{\sqrt{\frac{\pi n}{2}}}.
\end{aligned}$$

Here recall that

$$\left(1+\frac{x}{n}\right)^{an} \rightarrow e^{xa} \text{ as } n \rightarrow \infty.$$

For an even n it is shorter,

$$\begin{aligned}
& \mathbf{P}\left(X=\frac{n}{2}\right) \\
= \quad & \binom{n}{n/2} 2^{-n}=\frac{n!}{\left(\frac{n}{2}\right)!\left(\frac{n}{2}\right)!} 2^{-n} \\
\approx \quad & \frac{\sqrt{2\pi n}\left(\frac{n}{e}\right)^{n}}{\sqrt{2\pi\frac{n}{2}}\left(\frac{n}{2e}\right)^{\frac{n}{2}} \sqrt{2\pi\frac{n}{2}}\left(\frac{n}{2e}\right)^{\frac{n}{2}}} 2^{-n}=\frac{1}{\sqrt{\frac{\pi}{2} n}}.
\end{aligned}$$

5 hw5

1. A box contains b blue and r red balls (total number of balls in the box is $n = b + r$).

All balls are removed at random one by one and arranged in a row. Let X_i be the number of red balls between the $(i-1)$th and ith blue ball drawn, $i = 2, \ldots, b$; Let X_1 be the number of red balls until the first blue ball shows up, and X_{b+1} be the number of red balls after the last blue ball drawn. Consider the random vector $X = (X_1, \ldots X_{b+1})$. The range of X are all the vectors $(k_1, \ldots, k_{b+1})$ with nonnegative integer components $k_1, \ldots, k_{b+1}$ such that $k_1 + \ldots + k_{b+1} = r$. For such a vector $(k_1, \ldots, k_{b+1})$ with $k_i \geq 0$, and $k_1 + \ldots + k_{b+1} = r$, find

$$\mathbf{P}(X = (k_1, \ldots, k_{b+1})) = \mathbf{P}(X_1 = k_1, \ldots, X_{b+1} = k_{b+1}).$$

Note $X_i = 0$ means there are no red balls between the $(i-1)$th and ith blue balls if $i = 2, \ldots, b$; $X_1 = 0$ means the row starts with blue ball; $X_{b+1} = 0$ means the last ball is blue.

Hint. The computation is simple. Do not overthink. The vector $(k_1, \ldots, k_{b+1})$ with nonnegative integer components determines specifically (uniquely) r red ball and b blue ball "seats" (positions). For instance, if $n = 5, b = 2, r = 3$, and $(k_1, k_2, k_3) = (0, 2, 1)$, then five "seat" color arrangement is $BRRBR$, and $\mathbf{P}(X = (0, 2, 1)) = ?$

Answer. Total number of orderings of n balls is $n!$. The vector $(k_1, \ldots, k_{b+1})$ with nonnegative integer components determines specifically (uniquely) r red ball and b blue ball "seats" (positions). The red balls and blue balls have to be arranged in those "seats" (positions): by multiplication principle, there are $r!b!$ ways to do that . For example, with $n = 5, b = 2, r = 3$, and $(k_1, k_2, k_3) = (0, 2, 1)$, then five "seat" color arrangement is $BRRBR$;with $(k_1, k_2, k_3) = (1, 0, 2)$ the color arangement is RBBRR: in both cases,

$$\mathbf{P}(X = (0, 2, 1)) = \mathbf{P}(BRRBR) = \frac{2!3!}{5!}, \mathbf{P}(X = (1, 0, 2)) = \mathbf{P}(RBBRR) = \frac{2!3!}{5!} \text{ etc}$$

Thus

$$\mathbf{P}(X = (k_1, \ldots, k_{b+1})) = \frac{b!r!}{n!} = \frac{1}{\binom{n}{r}}.$$

With a little bit of "overthinking", and using multiplication principle, we could count the arrangements of balls for a given $(k_1, \ldots, k_{b+1})$ as $b!\binom{r}{k_1, \ldots, k_{b+1}} k_1! \ldots k_{b+1}! = r!b!$.

Comments. a) $X = (X_1, \ldots, X_{b+1})$ is discrete uniform random vector. It assumes values in the set of all $(k_1, \ldots, k_{b+1})$ vectors with nonnegative integer components whose sum $k_1 + \ldots + k_{b+1} = r$. Any value in that finite set is assumed with the same probability $b!r!/n!$.

b) The vector $(k_1, \ldots, k_{b+1})$ with nonnegative integer components determines specifically (uniquely) r red ball and b blue ball "seats" (positions): it determines uniquely the "word" of length n formed by b letters B and r letters R. There are $\binom{n}{r} = \binom{n}{b} = \binom{n}{r,b}$ such distinct words, and probability of a single "word" is

$$\frac{b!r!}{\binom{n}{r} r!b!} = \frac{1}{\binom{n}{r}}.$$

2. a) Given 1000 married couples, compute the probability that, in at least three of them, both husband and wife were born on the same day.

b) Given 300000 married couples, compute the probability that, in at least three of them, both husband and wife were born on April 18.

Use binomial and its Poisson approximation. Is approximation accurate?

Answer. First, $\mathbf{P}$ (husband and wife born 4/18) $= \frac{1}{365^2}$, $\mathbf{P}$ (husband and wife born the same day) $= 365 \cdot \frac{1}{365^2} = \frac{1}{365}$.

a) Y = # of couples with both born the same day among 1000 is Binomial(n, p), with $p = \frac{1}{365}, n = 1000$. Thus

$$\begin{aligned}
\mathbf{P}(X \geq 3) &= 1 - \mathbf{P}(X \leq 2) = 1 - \mathbf{P}(X = 0) - \mathbf{P}(X = 1) - \mathbf{P}(X = 2) \\
& \quad 1 - \left(\frac{364}{365}\right)^{1000} - 1000 \cdot \left(\frac{1}{365}\right)\left(\frac{364}{365}\right)^{999} - \binom{1000}{2}\left(\frac{1}{365}\right)^2\left(\frac{364}{365}\right)^{998} \\
&= 0.516\,3.
\end{aligned}$$

Also, using Poisson approximation, with $\lambda = np = \frac{1000}{365} = 2.\,739\,7$,

$$\begin{aligned}
\mathbf{P}(X \geq 3) &= 1 - \mathbf{P}(X \leq 2) \\
&= 1 - \mathbf{P}(X = 0) - \mathbf{P}(X = 1) - \mathbf{P}(X = 2) \\
&\approx 1 - e^{-\lambda}(1 + \lambda + \frac{\lambda^2}{2}) \\
&= 1 - e^{-\frac{1000}{365}}(1 + \frac{1000}{365} + \frac{\left(\frac{1000}{365}\right)^2}{2}) \\
&= 1 - \frac{39\,929}{5329} e^{-\frac{200}{73}} = 0.516\,06
\end{aligned}$$

Theoretical error estimate $np^2 = 1000\left(\frac{1}{365}\right)^2 = 7.\,506\,1 \times 10^{-3}$.

b) X = # of couples with both born 4/18 among 300000 is Binomial(n, p), with $p = \frac{1}{365^2}, n = 300000$. Then

$$\begin{aligned}
\mathbf{P}(X \geq 3) &= 1 - \mathbf{P}(X \leq 2) = 1 - \mathbf{P}(X = 0) - \mathbf{P}(X = 1) - \mathbf{P}(X = 2) \\
& \quad 1 - \left(1 - \frac{1}{365^2}\right)^{300000} - 300000 \cdot \left(\frac{1}{365^2}\right)\left(1 - \frac{1}{365^2}\right)^{299999} \\
& \quad -150000 \cdot 299999 \left(\frac{1}{365^2}\right)^2 \left(1 - \frac{1}{365^2}\right)^{299998}.
\end{aligned}$$

My calculator is unable to compute it.

Using Poisson approximation, with $\lambda = np = \frac{300000}{365^2} = 2.\,251\,8$, we have

$$\begin{aligned}
\mathbf{P}(X \geq 3) &= 1 - \mathbf{P}(X \leq 2) \\
&= 1 - \mathbf{P}(X = 0) - \mathbf{P}(X = 1) - \mathbf{P}(X = 2) \\
&\approx 1 - e^{-\lambda}(1 + \lambda + \frac{\lambda^2}{2}) \\
&= 1 - e^{-\frac{300000}{365^2}}(1 + \frac{300000}{365^2} + \frac{\left(\frac{300000}{365^2}\right)^2}{2}) \\
&= 0.391\,15.
\end{aligned}$$

Theoretical error estimate $np^2 = 300000 \left(\frac{1}{365^2}\right)^2 = 1.6902 \times 10^{-5}$ is smaller than in a).

3. a) Let X and Y have joint df F. Show that for any, $a < b$ and $c < d$,

$$\begin{aligned} & \mathbf{P}(a < X \le b, c < Y \le d) \\ = \; & F(b,d) - F(a,d) - F(b,c) + F(a,c). \end{aligned}$$

Find (in terms of F) the probability

$$\mathbf{P}(X = b, Y = d).$$

Comment. This equality implies that if F is a joint distribution function, then for any $a < b$ and $c < d$,

$$F(b,d) - F(a,d) - F(b,c) + F(a,c) \ge 0 \tag{5.1}$$

If F is twice continuously differentiable around the rectangle $(a,b] \times (c,d]$, using Taylor formula, we have

$$\begin{aligned} 0 & \le F(b,d) - F(a,d) - F(b,c) + F(a,c) \\ & = \int_0^1 \int_0^1 \frac{\partial^2 F(a + s(b-a), c + r(d-c))}{\partial x \partial y} drds \;\; (d-c)(b-a). \end{aligned}$$

Therefore,

$$\frac{\partial^2 F(a,c)}{\partial x \partial y} \ge 0 \tag{5.2}$$

and $\frac{\partial^2 F(x,y)}{\partial x \partial y} \ge 0$ in the rectangle above guaranties that (5.1) holds.

Answer. Let $a < b, c < d$. We could draw a picture. Alternatively,

$$\{a < X \le b, c < Y \le d\} = \{a < X \le b, Y \le d\} \setminus \{a < X \le b, Y \le c\}.$$

Since $\{a < X \le b, Y \le c\} \subseteq \{a < X \le b, Y \le d\}$,

$$\begin{aligned} & \mathbf{P}(a < X \le b, c < Y \le d) \\ = \; & \mathbf{P}(a < X \le b, Y \le d) - \mathbf{P}(a < X \le b, Y \le c). \end{aligned} \tag{5.3}$$

Similarly,

$$\begin{aligned} \mathbf{P}(a < X \le b, Y \le c) &= \mathbf{P}(X \le b, Y \le c) - \mathbf{P}(X \le a, Y \le c), \\ \mathbf{P}(a < X \le b, Y \le d) &= \mathbf{P}(X \le b, Y \le d) - \mathbf{P}(X \le a, Y \le d) \end{aligned}$$

Hence, by (5.3),

$$\begin{aligned} & \mathbf{P}(a < X \le b, c < Y \le d) \\ = \; & \mathbf{P}(a < X \le b, Y \le d) - \mathbf{P}(a < X \le b, Y \le c) \\ = \; & \mathbf{P}(X \le b, Y \le d) - \mathbf{P}(X \le a, Y \le d) \\ & - [\mathbf{P}(X \le b, Y \le c) - \mathbf{P}(X \le a, Y \le c)] \\ = \; & F(b,d) - F(a,d) - F(b,c) + F(a,c). \end{aligned}$$

By continuity of probability,

$$
\begin{aligned}
\mathbf{P}(X=b,Y=d) &= \lim_{n\to\infty}\mathbf{P}(b-1/n<X\le b, d-1/n<Y\le d)\\
&= \lim_{n\to\infty}\left[F(b,d)-F(b-1/n,d)-F(b,d-1/n)+F(b-1/n,d-1/n)\right]\\
&= F(b,d)-F(b-,d)-F(b,d-)+F(b-,d-).
\end{aligned}
$$

b) Is the function $F(x,y)=1-e^{-xy}, 0\le x,y<\infty$, the joint df of some pair of r.v.?

Comment. In order for F to be joint distribution function, besides the properties (a), (b), (c) of Lemma 5, p.39, the inequality (5.1) must hold as well, which translates into (5.2), and $\frac{\partial^2 F(x,y)}{\partial x\partial y}\ge 0$ if F is twice continuously differentiable around (x,y).

Answer. According to the comment above we must have $\frac{\partial^2 F(x,y)}{\partial x\partial y}\ge 0$ at any (x,y) where F is continuously differentiable. For $x,y>0$,

$$\frac{\partial^2 F(x,y)}{\partial x\partial y}=(1-xy)e^{-xy}<0 \text{ if } xy>1;$$

for instance if $x>1$ and $y>1$. Therefore F cannot be a joint density function.

Note that (a), (b), (c) of Lemma 5, p.39, hold for F.

4. Let $X_1,\ldots,X_n$ be identically distributed continuous random variables, i.e. they have the same df and pdf f. Assume that for any $-\infty<x_1,x_2,\ldots,x_n<\infty$,

$$\mathbf{P}(X_1\le x_1,X_2\le x_2,\ldots,X_n\le x_n)=\mathbf{P}(X_1\le x_1)\mathbf{P}(X_2\le x_2)\ldots\mathbf{P}(X_n\le x_n).$$

Such a collection of r.v.'s is called a random sample: we have n independent "observations" of a random variable X with df F and pdf f.

a) Show that $(X_1,\ldots,X_n)$ is jointly continuous with joint df

$$G(x_1,\ldots,x_n)=\int_{-\infty}^{x_1}\ldots\int_{-\infty}^{x_n} f(u_1)f(u_2)\ldots f(u_n)\,du_n\ldots du_1,$$

and joint pdf

$$g(x_1,\ldots,x_n)=f(x_1)f(x_2)\ldots f(x_n),(x_1,\ldots,x_n)\in\mathbf{R}^n.$$

Hint. Use independence to find joint df, then take mixed derivative in $x_1,\ldots,x_n$.

Answer. The joint df, by independence,

$$
\begin{aligned}
G(x_1,\ldots,x_n) &= \mathbf{P}(X_1\le x_1,\ldots,X_n\le x_n)=\mathbf{P}(X_1\le x_1)\ldots\mathbf{P}(X_n\le x_n)\\
&= \int_{-\infty}^{x_1} f(u_1)\,du_1\ldots\int_{-\infty}^{x_n} f(u_n)\,du_n\\
&= \int_{-\infty}^{x_1}\ldots\int_{-\infty}^{x_n} f(u_1)f(u_2)\ldots f(u_n)\,du_n\ldots du_1,(x_1,\ldots,x_n)\in\mathbf{R}^n.
\end{aligned}
$$

Then, by definition, $(X_1,\ldots,X_n)$ is continuous random vector with the joint pdf

$$g(x_1,\ldots,x_n)=f(x_1)f(x_2)\ldots f(x_n),(x_1,\ldots,x_n)\in\mathbf{R}^n.$$

b) Explain why with probability 1 we have $X_i\ne X_j$ for any $i\ne j$ (which means that with probability 1, the sample values $X_1(\omega),\ldots,X_n(\omega),\omega\in\Omega$, are all distinct)?

Hint. For any pair $i \neq j$, the pair (X_i, X_j) is jointly continuous.

Answer. According to part a), for any pair $i \neq j$, the pair (X_i, X_j) is jointly continuous with joint pdf

$$h(x_i, x_j) = f(x_i) f(x_j), (x_i, x_j) \in \mathbf{R}^2.$$

Hence, denoting $\Delta = \{(x_i, x_j) : x_i = x_j\}$ the line $x_i = x_j$ in x_i, x_j -plane, we have

$$\mathbf{P}(X_i = X_j) = \int\int_{x_i = x_j} f(x_i) f(x_j) \, dx_i dx_j = \int\int_{\Delta} f(x_i) f(x_j) \, dx_i dx_j = 0$$

because the double integral represents the volume above the line under the graph of $z = f(x_i) f(x_j)$.

5. Let X and Y be independent random variables taking values in the positive integers and having the same mass function $f(x) = 2^{-x}$ for $x \in \{1, 2, \ldots\}$, that is they are geometric with $p = 1/2$. Find their joint probability mass function and:

(a) $\mathbf{P}(\min\{X, Y\} \leq x)$. Hint. Find $\mathbf{P}(\min\{X, Y\} > x)$.

(b) $\mathbf{P}(Y > X)$; (c) $\mathbf{P}(X = Y)$, (d) $\mathbf{P}(X \geq kY)$, for a given positive integer k;

(e) $\mathbf{P}(X$ divides $Y)$. Hint. X divides Y means $Y = lX$ for some $l \in \{1, 2, \ldots\}$. Answer is a series.

Answer. Their joint probability mass function is

$$\begin{aligned} f(x, y) &= \mathbf{P}(X = x, Y = y) = \mathbf{P}(X = x) \mathbf{P}(Y = y) \\ &= 2^{-x} 2^{-y} \text{ if } x.y \in \{1, 2, \ldots\}, \end{aligned}$$

and $f(x, y) = 0$ otherwise.

(a) For $x \in \{1, 2, \ldots\}$

$$\begin{aligned} \mathbf{P}(\min\{X, Y\} \leq x) &= 1 - \mathbf{P}(\min\{X, Y\} > x) \\ &= 1 - \mathbf{P}(X > x) \mathbf{P}(Y > x) \\ &= 1 - \sum_{j=x+1}^{\infty} 2^{-j} \sum_{i=x+1}^{\infty} 2^{-i} = 1 - \left(\sum_{i=x+1}^{\infty} 2^{-i} \right)^2 \\ &= 1 - \left(\frac{2^{-(x+1)}}{1 - 2^{-1}} \right)^2 = 1 - 4 \cdot 2^{-2(x+1)} = 1 - 2^{-2x}. \end{aligned}$$

Finding (b)$\mathbf{P}(Y > X)$; (c) $\mathbf{P}(X = Y)$, (d) $\mathbf{P}(X \geq kY)$, for a given positive integer k;

Answer. (b) for $x, y \in \{1, 2, \ldots\}$,

$$\begin{aligned} \mathbf{P}(Y > X) &= \sum_{y > x} 2^{-x} 2^{-y} = \sum_{x=1}^{\infty} \sum_{y=x+1}^{\infty} 2^{-x} 2^{-y} = \sum_{x=1}^{\infty} 2^{-x} \sum_{y=x+1}^{\infty} 2^{-y} \\ &= \sum_{x=1}^{\infty} 2^{-x} \frac{2^{-(x+1)}}{1 - 1/2} = \sum_{x=1}^{\infty} 2^{-2x} = \frac{2^{-2}}{1 - 1/4} = 1/3. \end{aligned}$$

By symmetry of joint pmf, $\mathbf{P}(Y < X) = 1/3$ as well.

(c) for $x, y \in \{1, 2, \ldots\}$,

$$\begin{aligned}
\mathbf{P}(X = Y) &= \sum_{x=1}^{\infty} 2^{-x} 2^{-x} = \sum_{x=1}^{\infty} 2^{-2x} = \frac{2^{-2}}{1 - 1/4} = 1/3, \\
1 - \mathbf{P}(X \neq Y) &= 1 - \{\mathbf{P}(X < Y) + \mathbf{P}(X > Y)\} = 1/3.
\end{aligned}$$

Also, $\mathbf{P}(X = Y) = 1 - \{\mathbf{P}(X < Y) + \mathbf{P}(X > Y)\} = 1 - (1/3 + 1/3) = 1/3$.

(d) for $x, y \in \{1, 2, \ldots\}$,

$$\begin{aligned}
\mathbf{P}(X \geq kY) &= \sum_{x \geq ky} 2^{-x} 2^{-y} = \sum_{y=1}^{\infty} \sum_{x=ky}^{\infty} 2^{-x} 2^{-y} = \sum_{y=1}^{\infty} 2^{-y} \sum_{x=ky}^{\infty} 2^{-x} \\
&= \sum_{y=1}^{\infty} 2^{-y} \frac{2^{-ky}}{1 - 1/2} = 2 \sum_{y=1}^{\infty} 2^{-(k+1)y} = 2 \frac{2^{-(k+1)}}{1 - 2^{-(k+1)}} = \frac{2^{-k}}{1 - 2^{-(k+1)}} \\
&= \frac{2}{2^{k+1} - 1}.
\end{aligned}$$

(e) $\mathbf{P}(X \text{ divides } Y)$. Hint. X divides Y means $Y = kX$ for some $k \in \{1, 2, \ldots\}$. Answer is a series.

Answer. Since $\{X \text{ divides } Y\} = \cup_{k=1}^{\infty} \{Y = kX\}$,

$$\begin{aligned}
\mathbf{P}(X \text{ divides } Y) &= \sum_{k=1}^{\infty} \mathbf{P}(Y = kX) \\
&= \sum_{k=1}^{\infty} \sum_{y=kx} 2^{-x} 2^{-y} = \sum_{k=1}^{\infty} \sum_{x=1}^{\infty} 2^{-x} 2^{-kx} = \sum_{k=1}^{\infty} \sum_{x=1}^{\infty} 2^{-(k+1)x} \\
&= \sum_{k=1}^{\infty} \frac{2^{-(k+1)}}{1 - 2^{-(k+1)}} = \sum_{k=1}^{\infty} \frac{2^{-k}}{2 - 2^{-k}} = 0.6067.
\end{aligned}$$

6 hw6

1. (Continuation of #1 of hw5) A box contains b blue and r red balls (total number of balls in the box is $n = b + r$). All balls are removed at random one by one and arranged in a row. Let X_i be the number of red balls between the $(i-1)$th and ith blue ball drawn, $i = 2, \ldots, b$; Let X_1 be the number of red balls until the first blue ball shows up, and X_{b+1} be the number of red balls after the last blue ball drawn. In #5 of hw4, we found the pmf of X :

$$\begin{aligned} \mathbf{P}\left(X = (k_1, \ldots, k_{b+1})\right) &= \mathbf{P}\left(X_1 = k_1, \ldots, X_{b+1} = k_{b+1}\right) \\ &= f\left(k_1, \ldots, k_{b+1}\right) = \frac{b!r!}{n!} = \frac{1}{\binom{n}{r}} \end{aligned} \tag{6.4}$$

for any nonnegative integers $k_1, \ldots, k_{b+1}$ so that $k_1 + \ldots + k_{b+1} = r$.

a) Find $\mathbf{E}\left(X_1\right), \ldots, \mathbf{E}\left(X_{b+1}\right)$. Hint. The pmf f of X in (6.4) is symmetric in $(k_1, \ldots, k_{b+1})$: if we shuffle (rearrange) $k_1, \ldots, k_{b+1}$ the value of f does not change (it is still $f\left(k_1, \ldots, k_{b+1}\right)$). Therefore all the components $X_1, \ldots, X_{b+1}$ are identically distributed (they have the same pmf, that is all marginal pmfs are identical).

Answer. The pmf f of X in (6.4) is symmetric in $(k_1, \ldots, k_{b+1})$: if we shuffle (rearrange) $k_1, \ldots, k_{b+1}$ the value of f does not change (it is still $f\left(k_1, \ldots, k_{b+1}\right)$). Therefore all the components $X_1, \ldots, X_{b+1}$ are identically distributed. Therefore $\mathbf{E}\left(X_1\right) = \mathbf{E}\left(X_2\right) = \ldots = \mathbf{E}\left(X_{b+1}\right)$. On the other hand, $r = X_1 + \ldots + X_{b+1}$, and taking expectation of both sides,

$$r = \mathbf{E}\left(X_1\right) + \ldots + \mathbf{E}\left(X_{b+1}\right) = (b+1)\,\mathbf{E}\left(X_1\right),$$

and

$$\mathbf{E}\left(X_1\right) = \frac{r}{b+1} = \mathbf{E}\left(X_2\right) = \ldots = \mathbf{E}\left(X_{b+1}\right).$$

Comment on a). Why X_i are identically distributed? The range of every $X_i = \{0, 1, \ldots, r\}$. Then for instance, for $k \in \{0, 1, \ldots, r\}$, using symmetry and relabeling variable of summation,

$$\begin{aligned} f_{X_1}(k) &= \sum_{k+k_2+k_3+\ldots+k_{b+1}=r} f\left(k, k_2, k_3, \ldots, k_{b+1}\right) = \sum_{k+k_2+k_3+\ldots+k_{b+1}=r} f\left(k_2, k, k_3, \ldots, k_{b+1}\right) \\ & \sum_{k+k_1+k_3+\ldots+k_{b+1}=r} f\left(k_1, k, k_3, \ldots, k_{b+1}\right) = f_{X_2}(k) \text{ etc.} \end{aligned}$$

b) Let Y_i be the number of balls needed to be removed until the ith blue ball shows up, $i = 1, \ldots, b$. Find $\mathbf{E}\left(Y_i\right), i = 1, \ldots, b$. Hint. $Y_1 = X_1 + 1$.

Answer. By definition of X_i, we see that

$$\begin{aligned} Y_1 &= X_1 + 1, Y_2 = X_1 + X_2 + 2, \ldots, \\ Y_i &= X_1 + \ldots + X_i + i, i = 1, \ldots, b. \end{aligned}$$

So,

$$\mathbf{E}\left(Y_i\right) = i + \mathbf{E}\left(X_1 + \ldots + X_i\right) = i + i\frac{r}{b+1} = i\frac{r+b+1}{b+1} = i\frac{n+1}{b+1}, i = 1, \ldots, b.$$

c) Find the pmf of X_1 and Y_1. What are the pmf of $X_2, \ldots, X_{b+1}$?

Answer. We find pmf of Y_1 first. Range of $Y_1 = \{1, 2, \ldots, r+1\}$. For $k \in \{1, \ldots, r+1\}$, the total number of orderings of k balls out of n is $\binom{n}{k} k!$. Then we count the orderings of k balls so that the last ball is blue and $k-1$ before it are red. That number is $\binom{r}{k-1}(k-1)!b$. So,

$$\begin{aligned}
\mathbf{P}(Y_1 = k) &= \frac{\binom{r}{k-1}(k-1)!b}{\binom{n}{k}k!} = \frac{\binom{r}{k-1}b}{\binom{n}{k}k}, k = 1, \ldots, r+1, \\
\mathbf{P}(X_1 = j) &= \mathbf{P}(Y_1 = X_1 + 1 = j+1) = \frac{\binom{r}{j}b}{\binom{n}{j+1}(j+1)}, , j = 0, \ldots, r.
\end{aligned}$$

Comment on c). Alternatively, the same answer could be given the following way. There are $\binom{n}{r}$ distinct words of length n formed of b letters B and r letters R, and $\binom{n}{r}b!r! = n!$ distinct orderings of the balls.

Thinking about assigning "seats" in a row, there are $\binom{n}{r}$ distinct seats assignments, and $\binom{n}{r}b!r! = n!$ different orderings.

Then we count all the words of length n with first $k-1$ letters R and kth letter B: since we need to finish those words, it coincides with the number of words of length $n-k$ formed by $b-1$ letters B and remaining R's. So, we have $\binom{n-k}{b-1}b!r!$ orderings, and

$$\begin{aligned}
\mathbf{P}(Y_1 = k) &= \frac{\binom{n-k}{b-1}b!r!}{\binom{n}{r}b!r!} = \frac{\binom{n-k}{b-1}}{\binom{n}{r}}, k = 1, \ldots, r+1, \\
\mathbf{P}(X_1 = j) &= \mathbf{P}(Y_1 = X_1 + 1 = j+1) = \frac{\binom{n-1-j}{b-1}}{\binom{n}{r}}, j = 0, \ldots, r.
\end{aligned}$$

2. A biased coin is tossed n times with probability p of H. A run is a sequence of throws which result in the same outcome. For example, HHTHTTH contains 5 runs. Show that the expected number of runs is $1 + 2(n-1)pq$, where $q = 1 - p$. Find the variance of the number of runs. Hint: use indicators of A_j=" jth and $j+1$ -th outcomes are different".

Answer. Let $X_j = I_{A_j}$, A_j=" jth and $j+1$ -th outcomes are different". Then, denoting $q = 1-p$,

$$\mathbf{E}(X_j) = \mathbf{P}(A_j) = \mathbf{P}(H_j T_{j+1}) + \mathbf{P}(T_j H_{j+1}) = pq + pq = 2pq.$$

The total number of runs $X = 1 + \sum_{j=1}^{n-1} X_j$, and

$$\mathbf{E}(X) = 1 + \sum_{j=1}^{n-1} \mathbf{E}(X_j) = 1 + (n-1)2pq.$$

Using the variance covariance expansion, $Cov(X_i, X_j) = 0$ if $j > i+1$ by independence,

$$\begin{aligned}
\mathrm{Var}(X) &= \mathrm{Var}\left(\sum_{j=1}^{n-1} X_j\right) = \sum_{j=1}^{n-1} \mathrm{Var}(X_j) + 2\sum_{i<j} \mathrm{Cov}(X_i, X_j) \\
&= \sum_{j=1}^{n-1} \left[2pq - (2pq)^2\right] + 2\sum_{i=1}^{n-2} \mathrm{Cov}(X_i, X_{i+1}).
\end{aligned}$$

Now,

$$\begin{aligned}
\operatorname{Cov}(X_i, X_{i+1}) &= \mathbf{P}(A_i \cap A_{i+1}) - \mathbf{P}(A_i)\,\mathbf{P}(A_j) \\
&= \mathbf{P}(H_i T_{i+1} H_{i+2}) + \mathbf{P}(T_i H_{i+1} T_{i+2}) - (2pq)^2 \\
&= p^2 q + pq^2 - (2pq)^2 = pq - 4p^2 q^2.
\end{aligned}$$

Thus

$$\begin{aligned}
&\operatorname{Var}(X) \\
={}& (n-1)\left(2pq - 4p^2q^2\right) + 2(n-2)\left(pq - 4p^2q^2\right) \\
={}& (2n-3)\,2pq + (5-3n)\,4p^2q^2 = 2pq\,[2n - 3 + (5-3n)\,2pq]\,.
\end{aligned}$$

3. A building has 10 floors above the street level. 20 of people enter at street level and board an elevator. They choose floors independently, equally likely at random. Let T be the number of stops the elevator must make.

Find $\mathbf{E}(T)$and Var(T). Hint. Use the indicators of the event that ith floor selected. Note "ith floor selected" = "at least one person selects floor i". What is the probability that no one selects the ith floor?

Answer. Let $A_i =$ "ith floor selected", $i = 1, \ldots, 10$. We have

$$T = \sum_{i=1}^{10} I_{A_i}, \mathbf{E}(T) = \sum_{i=1}^{10} \mathbf{E}\left(I_{A_i}\right) = \sum_{i=1}^{10} \mathbf{P}(A_i)\,.$$

By independence of individual floor choices of different people,

$$\mathbf{P}(A_i) = 1 - \mathbf{P}(\text{nobody selects floor } i) = 1 - \left(\frac{9}{10}\right)^{20} = 0.878\,42,$$

and

$$\mu := \mathbf{E}(T) = 10\left(1 - \left(\frac{9}{10}\right)^{20}\right) = 8.78\,42.$$

Since

$$\operatorname{Var}(T) = E\left(T^2\right) - \mu^2,$$

we find

$$T^2 = \sum_{i=1}^{10} I_{A_i} + 2\sum_{i<j} I_{A_i A_j}, \mathbf{E}\left(T^2\right) = \sum_{i=1}^{10} \mathbf{P}(A_i) + 2\sum_{i<j} \mathbf{P}\left(A_i A_j\right) = \mu + 2\sum_{i<j} \mathbf{P}\left(A_i A_j\right).$$

For $i < j$,

$$\begin{aligned}
\mathbf{P}(A_i A_j) &= 1 - \mathbf{P}\left(A_i^c \cup A_j^c\right) = 1 - \mathbf{P}(A_i^c) - \mathbf{P}(A_j^c) + \mathbf{P}(A_i^c A_j^c) \\
&= 1 - 2\cdot\frac{9^{20}}{10^{20}} + \frac{8^{20}}{10^{20}} =: a, \text{ and} \\
\mathbf{E}\left(T^2\right) &= \mu + 2\cdot\frac{10\cdot 9}{2} a = \mu + 90a
\end{aligned}$$

Hence

$$\operatorname{Var}(T) = E\left(T^2\right) - \mu^2 = \mu + 90a - \mu^2 = 0.775\,31.$$

4. *Coupons*. Every package of some intrinsically dull commodity includes a small and exciting plastic object. There are n different types of object, and each package is equally likely to contain any given type. You buy one package each day.

(a) Let X_j be number of days which elapse between the acquisitions of the jth new type of object and the $(j+1)$th new type. Recognize the distribution of X_j and find $\mathbf{E}\left(X_j\right)$. Hint. Since we have already j objects there are $n-j$ new left. After the acquisition of the jth new type, what is the probability that the next package will contain nothing new?

Answer. After the acquisition of the jth new type, we have j objects and there are $n-j$ new left. Then the next package will contain nothing new with probability $\frac{j}{n}$ and something new with probability $1-\frac{j}{n} = \frac{n-j}{n}$. Since the days (packages) are independent, X_j is geometric with $p = \frac{n-j}{n}$, and we know

$$\mathbf{E}\left(X_j\right) = \frac{1}{p} = \frac{n}{n-j}, \; j = 1, \ldots, n-1.$$

(b) Let X be the number of days needed for you to have a full set of objects. Find $\mathbf{E}(X)$. Express X using X_j.

Answer. Since $X = 1 + X_1 + \ldots + X_{n-1}$, we have (changing the variable of summation, $k = n-j$)

$$\begin{aligned}
\mathbf{E}(X) &= 1 + \sum_{j=1}^{n-1} \mathbf{E}\left(X_j\right) = 1 + \sum_{j=1}^{n-1} \frac{n}{n-j} = \sum_{j=0}^{n-1} \frac{n}{n-j} \\
&= n\left(1 + \frac{1}{2} + \ldots + \frac{1}{n-1} + \frac{1}{n}\right).
\end{aligned}$$

(c) Let A_i be the event that none of the first k packages contain the ith object. Find $\mathbf{P}(A_1 \cup A_2 \cup A_3 \cup A_4)$.

Answer. By inclusion/exclusion and independence of packages,

$$\begin{aligned}
\mathbf{P}(A_1 \cup A_2 \cup A_3 \cup A_4) &= \sum_{i=1}^{4} \mathbf{P}(A_i) - \sum_{i<j} \mathbf{P}\left(A_i A_j\right) + \binom{4}{3} \sum_{i<j<k} \mathbf{P}\left(A_i A_j A_k\right) - \mathbf{P}(A_1 A_2 A_3 A_4) \\
&= 4\left(1-\frac{1}{n}\right)^k - \binom{4}{2}\left(1-\frac{2}{n}\right)^k + \binom{4}{3}\left(1-\frac{3}{n}\right)^k - \left(1-\frac{4}{n}\right)^k \\
&= 4\left(1-\frac{1}{n}\right)^k - 6\left(1-\frac{2}{n}\right)^k + 4\left(1-\frac{3}{n}\right)^k - \left(1-\frac{4}{n}\right)^k.
\end{aligned}$$

5. You roll a fair die repeatedly. If it shows 1, you must stop, but you may choose to stop at any prior time. Your score is the number shown by the die on the final roll. Consider the following strategy $S(k)$: stop the first time that the die shows k or greater, $k = 4, 6$.

a) What is the probability that "k or greater" ($k = 4, 6$) shows up before 1? What is the probability that 1 shows up before "k or greater" ($k = 4, 6$)?

Hint. Consider A_k= "k or greater shows up before 1", and use first step analysis to write an equation for $\mathbf{P}(A_k)$.

Answer. Let D_1 ="1 in the first roll", D_k = "$\geq k$ in the first roll", $\bar{D}_k$ = "neither 1 nor $\geq k$ in the first roll". If a fair die is rolled, for $k = 2, 3, 4, 5, 6$,

$$\begin{aligned}\mathbf{P}(D_4) &= \frac{3}{6} = \frac{1}{2}, \mathbf{P}(D_1) = \mathbf{P}(D_6) = \frac{1}{6},\\ \mathbf{P}(\bar{D}_6) &= \frac{4}{6} = \frac{2}{3}, \mathbf{P}(\bar{D}_4) = \frac{2}{6} = \frac{1}{3}.\end{aligned}$$

By total probability law,

$$\begin{aligned}\mathbf{P}(A_k) &= \mathbf{P}(A_k|D_1)\mathbf{P}(D_1) + \mathbf{P}(A_k|D_k)\mathbf{P}(D_k) + \mathbf{P}(A_k|\bar{D}_k)\mathbf{P}(\bar{D}_k)\\ &= \mathbf{P}(A_k|D_k)\mathbf{P}(D_k) + \mathbf{P}(A_k|\bar{D}_k)\mathbf{P}(\bar{D}_k),\end{aligned}$$

and $\mathbf{P}(A_k|D_k) = 1, \mathbf{P}(A_k|\bar{D}_k) = \mathbf{P}(A_k)$. So, we get an equation for $\mathbf{P}\left(A^k\right)$:

$$\mathbf{P}(A_k) = \mathbf{P}(D_k) + \mathbf{P}(A_k)\mathbf{P}(\bar{D}_k),$$

and

$$\mathbf{P}(A_k) = \frac{\mathbf{P}(D_k)}{1 - \mathbf{P}(\bar{D}_k)}.$$

We find

$$\mathbf{P}(A_6) = \frac{\frac{1}{6}}{1 - \frac{2}{3}} = \frac{1}{2}, \mathbf{P}(A_4) = \frac{\frac{1}{2}}{1 - \frac{1}{3}} = \frac{3}{4}.$$

b) Let X_k be the score when the strategy $S(k)$ was used. Find $\mathbf{E}(X_k), k = 4, 6$. Which strategy yields the higher expected score?

Answer. Range of $X_4 = \{1, 4, 5, 6\}$, Range of $X_6 = \{1, 6\}$. X_4 value is determined by the score falling for the first time into $\{1, 4, 5, 6\}$ which happens with probability 1. Since the die is fair, all 4 values are equally likely,

$$\mathbf{P}(X_4 = 1) = \mathbf{P}(X_4 = 4) = \mathbf{P}(X_4 = 5) = \mathbf{P}(X_4 = 6) = \frac{1}{4}.$$

Hence

$$\mathbf{E}(X_4) = \frac{1}{4}(1 + 4 + 5 + 6) = 4.$$

Similarly,

$$\mathbf{P}(X_6 = 1) = \mathbf{P}(X_6 = 6) = \frac{1}{2},$$

and

$$\mathbf{E}(X_6) = \frac{1}{2}(1 + 6) = 3.5.$$

The strategy S_4 is better.

2nd answer.

$$\mathbf{E}(X_6) = 1 \cdot \mathbf{P}(1 \text{ before } 6) + 6 \cdot \mathbf{P}(6 \text{ before } 1) = \frac{1}{2} + \frac{6}{2} = 3.5,$$

and

$$\begin{aligned}
\mathbf{E}(X_4) &= \mathbf{E}(X_4|A_4)\,\mathbf{P}(A_4) + \mathbf{E}\left(X_4|A_4^c\right)\mathbf{P}\left(A_4^c\right) \\
&= \left(4 \cdot \frac{1}{3} + 5 \cdot \frac{1}{3} + 6 \cdot \frac{1}{3}\right)\frac{3}{4} + 1 \cdot \frac{1}{4} = 4
\end{aligned}$$

Comment. We could use series to compute the pmf of X_4. For instance, $\{X_4 = 5\} = \cup_{l=1}^{\infty}\{X_4 = 5 \text{ in the } l\text{th roll}\}$,

$$\mathbf{P}(X_4 = 5) = \sum_{l=1}^{\infty}\mathbf{P}(X_4 = 5 \text{ in the } l\text{th roll}) = \sum_{l=1}^{\infty}\left(\frac{2}{6}\right)^{l-1}\frac{1}{6} = \frac{1}{4},$$

because "$X_4 = 5$ in the lth roll" = " 2 or 3 in the first $l-1$ roll and 5 in the lth roll" etc.

7 hw7

1. An urn contains n balls numbered 1, 2, ..., n. We remove k balls at random without replacement, and add up their numbers. Find the mean and an expression for the variance of the total. Recall

$$1+2+\ldots+n=\frac{n(n+1)}{2}.$$

Answer. Let A_j be the event that the ball numbered j is among k removed, $j=1,\ldots,n$. Then the total $X=\sum_{j=1}^{n} jI_{A_j}$, and$\mathbf{E}(X)=\sum_{j=1}^{n} j\,\mathbf{E}\left(I_{A_j}\right)=\sum_{j=1}^{n} j\,\mathbf{P}\left(A_j\right)=\sum_{j=1}^{n} j\,\mathbf{P}\left(A_j\right)$.The sample space Ω are all different groups of k out of n: $\#\Omega=\binom{n}{k}$; A_j consists of all groups of k that contain the ball numbered j: $\#A_j=\binom{n-1}{k-1}$. Hence $P\left(A_j\right)=\frac{\binom{n-1}{k-1}}{\binom{n}{k}}=\frac{k}{n}$, and

$$m=\mathbf{E}(X)=\sum_{j=1}^{n} j\cdot\frac{k}{n}=k\frac{n(n+1)}{2n}=k\cdot\frac{n+1}{2}.$$

Now, $\mathrm{Var}(X)=\mathbf{E}\left(X^2\right)-m^2$, and

$$\mathbf{E}\left(X^2\right)=\mathbf{E}\left(\sum_{j=1}^{n} j^2 I_{A_j}\right)+2\sum_{i<j} ij\,\mathbf{E}\left(I_{A_i A_j}\right)=\sum_{j=1}^{n} j^2\mathbf{P}\left(A_j\right)+2\sum_{i<j} ij\,\mathbf{P}\left(A_i A_j\right)$$

We find

$$\mathbf{P}\left(A_i A_j\right)=\frac{\binom{n-2}{k-2}}{\binom{n}{k}}=\frac{(n-2)!k!(n-k)!}{(k-2)!(n-k)!n!}=\frac{k(k-1)}{n(n-1)},$$

and

$$\begin{aligned}
\mathrm{Var}(X) &= \mathbf{E}\left(X^2\right)-m^2=\frac{k}{n}\sum_{j=1}^{n} j^2-m^2+2\frac{k(k-1)}{n(n-1)}\sum_{i<j} ij \\
&= \frac{k}{n}\sum_{j=1}^{n} j^2-m^2+2\frac{k(k-1)}{n(n-1)}\sum_{i<j} ij
\end{aligned}$$

That is an expression for $\mathrm{Var}(X)$.

Comment. If we want to simplify, then

$$\begin{aligned}
\sum_{j=1}^{n} j^2 &= \frac{n(n+1)(2n+1)}{6}, \\
2\sum_{i<j} ij &= \sum_{i\neq j} ij=\sum_{i,j} ij-\sum_{i=1}^{n} i^2 \\
&= \left(\sum_{i=1}^{n} i\right)^2-\sum_{i=1}^{n} i^2=\frac{n^2(n+1)^2}{4}-\frac{n(n+1)(2n+1)}{6} \\
&= n(n+1)\left(\frac{n(n+1)}{4}-\frac{n}{3}-\frac{1}{6}\right)=\frac{n(n+1)(n-1)\left(n+\frac{2}{3}\right)}{4}
\end{aligned}$$

and

$$\begin{aligned}
\operatorname{var}(X) &= \frac{k}{n}\sum_{j=1}^{n} j^2 - m^2 + 2\frac{k(k-1)}{n(n-1)}\sum_{i<j} ij \\
&= \frac{k}{n}\frac{n(n+1)(2n+1)}{6} + \frac{k(k-1)}{n(n-1)}\frac{n(n+1)(n-1)\left(n+\frac{2}{3}\right)}{4} - m^2 \\
&= k\frac{(n+1)(2n+1)}{6} + k(k-1)\frac{(n+1)\left(n+\frac{2}{3}\right)}{4} - k^2\cdot\frac{(n+1)^2}{4} \\
&= k(n+1)\left[\frac{2n+1}{6} + (k-1)\frac{n+\frac{2}{3}}{4} - k\frac{n+1}{4}\right] \\
&= \frac{k(n-k)(n+1)}{12}.
\end{aligned}$$

2. a) Let

$$Y = g(X) + W,$$

and W and X are independent discrete r.v. Show that

$$\mathbf{P}(Y = y|X = x) = \mathbf{P}(g(x) + W = y).$$

Answer. By definition,

$$\begin{aligned}
\mathbf{P}(Y = y|X = x) &= \frac{\mathbf{P}(Y = y, X = x)}{\mathbf{P}(X = x)} = \frac{\mathbf{P}(g(X) + W = y, X = x)}{\mathbf{P}(X = x)} \\
&= \frac{\mathbf{P}(g(x) + W = y, X = x)}{\mathbf{P}(X = x)} = \frac{\mathbf{P}(g(x) + W = y)\,\mathbf{P}(X = x)}{\mathbf{P}(X = x)} \\
&= \mathbf{P}(g(x) + W = y).
\end{aligned}$$

b) Let

$$X = Y + U,$$

and Y and U are independent. Assume the pmf $f_Y(y)$ and $f_U(u)$ are known. Find the joint pmf $f(x, y)$ of X and Y, and

$$\mathbf{P}(Y = y|X = x)$$

in terms of f_U and f_Y.

Answer. The joint pdf

$$\begin{aligned}
f(x, y) &= \mathbf{P}(Y = y, Y + U = x) = \mathbf{P}(Y = y, y + U = x) \\
&= \mathbf{P}(Y = y, U = x - y) = f_Y(y)\, f_U(x - y),
\end{aligned}$$

and

$$\begin{aligned}
f_X(x) &= \mathbf{P}(X = x) = \sum_y f(x, y) = \sum_y f_Y(y)\, f_U(x - y) \\
&= \sum_r f_Y(r)\, f_U(x - r).
\end{aligned}$$

Hence

$$
\begin{aligned}
& \mathbf{P}(Y=y|X=x) \\
= \; & \frac{f(x,y)}{f_X(x)} = \frac{f_Y(y)\, f_U(x-y)}{\sum_r f_Y(r)\, f_U(x-r)}
\end{aligned}
$$

3. A factory has produced n robots, each of which is faulty with probability p. To each robot a test is applied which passes all good robots and detects the fault (if present) with probability δ. Let X be the number of faulty robots, and Y the number detected as faulty. Assume the usual independence.

(a) What is the probability that a robot passed is in fact faulty?

Answer. By definition,

$$
\begin{aligned}
\mathbf{P}(\text{faulty}|\text{passed}) &= \frac{\mathbf{P}(\text{faulty and passed})}{\mathbf{P}(\text{passed})} \\
&= \frac{\mathbf{P}(\text{passed}|\text{faulty})\,\mathbf{P}(\text{faulty})}{\mathbf{P}(\text{passed}|\text{faulty})\,\mathbf{P}(\text{faulty}) + \mathbf{P}(\text{passed}|\text{good})\,\mathbf{P}(\text{good})} \\
&= \frac{(1-\delta)\,p}{(1-\delta)\,p + (1)(1-p)} = \frac{(1-\delta)\,p}{1-\delta p}.
\end{aligned}
$$

(b) Let Z be the number of passed faulty robots. Given $Y=k$, what is the distribution of Z? What is $\mathbf{E}(Z|Y)$?

Answer. Given $Y=k$, $Z \sim \text{bin}\left(n-k, \frac{(1-\delta)p}{1-\delta p}\right)$, and for $0 \le k \le n$,

$$
\mathbf{E}(Z|Y=k) = (n-k)\,\frac{(1-\delta)\,p}{1-\delta p}.
$$

Hence

$$
\mathbf{E}(Z|Y) = (n-Y)\,\frac{(1-\delta)\,p}{1-\delta p}.
$$

(c) Show that

$$
\mathbf{E}(X|Y) = \frac{np(1-\delta) + (1-p)\,Y}{1-p\delta}.
$$

Answer. We have $X = Y + Z$, and, since $\mathbf{E}(Y|Y) = Y$,

$$
\begin{aligned}
\mathbf{E}(X|Y) &= \mathbf{E}(Y|Y) + \mathbf{E}(Z|Y) = Y + \mathbf{E}(Z|Y) \\
&= Y + (n-Y)\,\frac{(1-\delta)\,p}{1-\delta p} = \frac{np(1-\delta) + (1-p)\,Y}{1-p\delta}.
\end{aligned}
$$

Comment. $\mathbf{E}(X|Y)$ found is the best mean square estimate of the total number of faulty robots based on the number of detected faulty robots.

4. a) We define the conditional variance, $\text{Var}(Y|X)$, as a random variable

$$
\text{Var}(Y|X) = \mathbf{E}\left(Y^2|X\right) - (\mathbf{E}(Y|X))^2.
$$

Show that

$$
\text{Var}(Y) = \mathbf{E}(\text{Var}(Y|X)) + \text{Var}(\mathbf{E}(Y|X)).
$$

Answer. By Theorem (4), p. 67,

$$\begin{aligned}\mathbf{E}\left(\mathrm{Var}\left(Y|X\right)\right) &= \mathbf{E}\left[\mathbf{E}\left(Y^2|X\right)\right]-\mathbf{E}\left[\left(\mathbf{E}\left(Y|X\right)\right)^2\right]\\ &= \mathbf{E}\left(Y^2\right)-\mathbf{E}\left[\left(\mathbf{E}\left(Y|X\right)\right)^2\right].\end{aligned}$$

On the other hand, $\mathbf{E}\left[\mathbf{E}\left(Y|X\right)\right]=\mathbf{E}\left(Y\right)$, and

$$\mathrm{Var}\left(\mathbf{E}\left(Y|X\right)\right)=\mathbf{E}\left[\left(\mathbf{E}\left(Y|X\right)\right)^2\right]-\left(\mathbf{E}\left(Y\right)\right)^2.$$

Thus

$$\mathbf{E}\left(\mathrm{Var}\left(Y|X\right)\right)+\mathrm{Var}\left(\mathbf{E}\left(Y|X\right)\right)=\mathrm{Var}\left(Y\right).$$

b) (application of a)) We have a coin that shows heads with probability p. We tossed it repeatedly and counted how many tosses were needed for the first heads to show up. If that number is X, we roll the fair die X times. Let Y be the total score in X rolls of the die.

Find $\mathbf{E}(Y|X=n)$, $\mathbf{E}(Y|X)$ and $\mathbf{E}(Y)$. Find $\mathrm{Var}(Y|X=n)$,$\mathrm{Var}(Y|X)$, $\mathrm{Var}(Y)$. Hint. $Y=\sum_{i=1}^{X}Y_i$, where Y_i is the score in the ith roll.

Answer. Note that X is geometric(p) : $\mathbf{E}\left(X\right)=\frac{1}{p}$,$\mathrm{Var}(X)=q/p^2$. We model Y as $Y=\sum_{i=1}^{X}Y_i$, where Y_i are independent die scores independent of X. Given $X=n$,

$$Y=\sum_{i=1}^{n}Y_i\,.$$

We find

$$\begin{aligned}\mathbf{E}\left(Y_i\right) &= \mu=\frac{1+2+3+4+5+6}{6}=\frac{7}{2},\\ \mathrm{Var}\left(Y_i\right) &= \sigma^2=\sum_{k=1}^{6}k^2\cdot\frac{1}{6}-\mu^2=\frac{35}{12}.\end{aligned}$$

Hence

$$\begin{aligned}\mathbf{E}(Y|X &= n)=\mathbf{E}\left(\sum_{i=1}^{n}Y_i\right)=\sum_{i=1}^{n}\mathbf{E}\left(Y_i\right)=n\mu,\\ \mathrm{Var}\left(Y|X=n\right) &= \mathrm{Var}\left(\sum_{i=1}^{n}Y_i\right)=\sum_{i=1}^{n}\mathrm{Var}\left(Y_i\right)=n\sigma^2,\end{aligned}$$

and $\mathbf{E}\left(Y|X\right)=\mu X$,$\mathrm{Var}(Y|X)=\sigma^2X$. So,

$$\mathbf{E}\left(Y\right)=\mathbf{E}\left[\mathbf{E}\left(Y|X\right)\right]=\mathbf{E}\left(\mu X\right)=\mu\mathbf{E}\left(X\right)=\frac{\mu}{p}.$$

By part a),

$$\begin{aligned}\mathrm{Var}\left(Y\right) &= \mathbf{E}\left[\mathrm{Var}\left(Y|X\right)\right]+\mathrm{Var}\left(\mathbf{E}\left(Y|X\right)\right)=\mathbf{E}\left(\sigma^2X\right)+\mathrm{Var}\left(\mu X\right)\\ &= \sigma^2\mathbf{E}\left(X\right)+\mu^2\mathrm{Var}\left(X\right)=\frac{\sigma^2}{p}+\mu^2\cdot\frac{q}{p^2}.\end{aligned}$$

5. 51 passengers bought tickets on a 51-seat carriage. One seat was reserved for each passenger. The first 50 passengers took the seats at random so that all 51! possible seating arrangements (with one empty seat) are equally likely. The last passenger insisted on taking the assigned seat. If that seat is occupied, then the passenger in that seat has to move to the corresponding assigned seat, and so on.

a) Find the probability that the last passengers seat is not empty.

Answer. Consider the events $A =$ "last passengers seat is empty" and $A^c =$ "last passengers seat is taken". In an n seat carriage,

$$\mathbf{P}(A) = \frac{(n-1)!}{n!} = \frac{1}{n}, \mathbf{P}\left(A^c\right) = 1 - \frac{1}{n} = \frac{n-1}{n}.$$

b) Compute the expected value of the number M of passengers who have to change their seats.

Hint. One way could be to denote E_n the expected value of the number of passengers who have to change their seats in n seat plane (instead of 51), and write a recursion for E_n in terms of E_{n-1}, and going "down" to get a general formula for E_n without knowing the distribution of $M = M_n$. You can obtain recursion by conditioning with respect to whether the last passengers seat is empty or not (1st step analysis). Another way is by simply finding the pmf of $M = M_n$ explicitly: directly or writing a recursion for $p_{n,k}$, probability that k passengers were moved in n seat plane, in terms of $p_{n-1,k-1}$.

1st answer. Let M_n be the number of passengers who have to change their seats in n seats plane. Using conditioning,

$$E_n = \mathbf{E}(M_n) = \mathbf{E}(M_n|A)\,\mathbf{P}(A) + \mathbf{E}\left(M_n|A^c\right)\mathbf{P}\left(A^c\right).$$

Given A, we have $M_n = 0$, and $\mathbf{E}(M_n|A) = 0$. Given A^c, the last passenger displaces one person and takes the assigned seat. The displaced passenger becomes like the last passenger: he faces $n-1$ remaining seats with one seat empty and wants to take the seat assigned to him. That is, given A^c, we have $M_n = 1 + M_{n-1}$ and$\mathbf{E}(M_n|A^c) = 1 + \mathbf{E}(M_{n-1}) = 1 + E_{n-1}$. So, we get the recursion

$$E_n = (1 + E_{n-1})\frac{n-1}{n} = \frac{n-1}{n} + E_{n-1}\frac{n-1}{n}.$$

Going "down" we get (note $E_2 = (1 + E_1)\frac{1}{2} = \frac{1}{2}, E_1 = 0$)

$$\begin{aligned}
E_n &= \frac{n-1}{n} + E_{n-1}\frac{n-1}{n} \\
&= \frac{n-1}{n} + (1 + E_{n-2})\frac{n-2}{n-1}\frac{n-1}{n} = \frac{n-1}{n} + (1 + E_{n-2})\frac{n-2}{n} \\
&= \frac{n-1}{n} + \frac{n-2}{n} + E_{n-2}\frac{n-2}{n} = ... \\
&= \frac{n-1}{n} + \frac{n-2}{n} + \ldots + \frac{2}{n} + E_2\frac{2}{n} = \frac{n-1}{n} + \frac{n-2}{n} + \ldots + \frac{2}{n} + \frac{1}{n} \\
&= \frac{1 + 2 + \ldots + (n-1)}{n} = \frac{n(n-1)}{2n} = \frac{n-1}{2}.
\end{aligned}$$

So, $E_{51} = 50/2 = 25$.

2nd answer. The range of M_n is $\{0, 1, \ldots, n-1\}$. Now

$$\#\{M_n = k\} = \binom{n-1}{k} k!\,(n-1-k)! = \frac{(n-1)!}{(n-k-1)!}\,(n-1-k)! = (n-1)!,$$

and

$$\mathbf{P}(M_n = k) = \frac{(n-1)!}{n!} = \frac{1}{n}, k = 0, 1, \ldots, n-1.$$

That is, M_n is discrete uniform r.v. in $\{0, 1, \ldots, n-1\}$, and

$$\mathbf{E}(M_n) = \frac{0 + 1 + \ldots + (n-1)}{n} = \frac{n(n-1)}{2n} = \frac{n-1}{2}.$$

3rd answer. Let $A_{n,k} =$ "k passengers were moved in n seat plane", $p_{n,k} = \mathbf{P}(A_{n,k})$. Let $A =$ "last passengers seat is empty" and $A^c =$ "last passengers seat is taken". Then

$$\begin{aligned} p_{n,k} &= \mathbf{P}(A_{n,k}) = \mathbf{P}(A_{n,k}|A)\,\mathbf{P}(A) + \mathbf{P}(A_{n,k}|A^c)\,\mathbf{P}(A^c) \\ &= \mathbf{P}(A_{n,k}|A^c)\,\mathbf{P}(A^c) = p_{n-1,k-1}\frac{n-1}{n} = \frac{n-1}{n}p_{n-1,k-1}. \end{aligned}$$

For any n, $p_{n,0} = \frac{1}{n}$, and going down,

$$\begin{aligned} p_{n,k} &= \frac{n-1}{n}p_{n-1,k-1} = \frac{n-1}{n}\frac{n-2}{n-1}p_{n-2,k-2} = \frac{n-2}{n}p_{n-2,k-2} \\ &= \ldots = \frac{n-k}{n}p_{n-k,0} = \frac{n-k}{n}\frac{1}{n-k} = \frac{1}{n}, 1 \le k \le n-1. \end{aligned}$$

Hence $\mathbf{P}(M_n = k) = p_{n,k} = 1/n$ for any $0 \le k \le n-1$. That is M_n is uniform in $\{0, 1, \ldots, n-1\}$, and we finish like in the 2nd answer.

8 hw8

1. Consider a simple random walk on the set $\{0, 1, \ldots, N\}$ in which each step is to the right with probability p or to the left with probability $q = 1 - p$. Absorbing barriers are placed at 0 and N. Let $S_0 = k, 0 < k < N$.

Show that the number X of positive steps of the walk before absorption satisfies

$$\mathbf{E}(X) = \frac{1}{2}\left[D_k - k + N(1 - p_k)\right],$$

where D_k is the mean number of steps until absorption and p_k is the probability of absorption at 0 (the gambler ruin probability). Hint. If Z_k is the number of steps until absorption, and Y is the number of negative steps until absorption, then $Z_k = X + Y$, $D_k = \mathbf{E}(Z_k) = \mathbf{E}(X + Y)$. What can you say about $k + X - Y$? How many values it takes?

Answer. If Z_k is the number of steps until absorption, and Y is the number of negative steps until absorption, then $Z_k = X + Y$, and

$$D_k = \mathbf{E}(Z_k) = \mathbf{E}(X + Y) = \mathbf{E}(X) + \mathbf{E}(Y).$$

Since $k + X - Y$ is the position of the walk at the absorption moment it takes two values:

$$k + X - Y = \begin{cases} 0 & \text{if absorbed at } 0 \\ N & \text{if absorbed at } N. \end{cases}$$

Hence

$$\mathbf{E}(k + X - Y) = k + \mathbf{E}(X) - \mathbf{E}(Y) = 0 \cdot p_k + N(1 - p_k) = N(1 - p_k).$$

Thus, adding the equations,

$$k + 2\mathbf{E}(X) = D_k + N(1 - p_k),$$

and

$$\mathbf{E}(X) = \frac{1}{2}\left[D_k - k + N(1 - p_k)\right].$$

2. Consider a symmetric simple random walk S_n with $S_0 = 0, p = q = 1/2$. Let $\tau_0 = \min\{n \geq 1 : S_n = 0\}$ be the time of the first return of the walk to its starting point.

Show that

$$\mathbf{P}(\tau_0 = 2n) = \frac{1}{2n - 1}\binom{2n}{n}2^{-2n} = \frac{1}{2n - 1}\mathbf{P}(S_{2n} = 0).$$

Hint. Recall we found and computed $\mathbf{P}(\tau_0 > 2n) = \mathbf{P}(S_{2n} = 0)$.

Answer. Since the τ_0 values are even integers,

$$\begin{aligned}
\mathbf{P}(\tau_0 = 2n) &= \mathbf{P}(2n - 2 < \tau_0 \leq 2n) = \mathbf{P}(\tau_0 > 2n - 2) - \mathbf{P}(\tau_0 > 2n) \\
&= \binom{2n - 2}{n - 1}2^{-(2n-2)} - \binom{2n}{n}2^{-2n} = \frac{(2n - 2)!4n}{(n - 1)!n!}2^{-2n} - \binom{2n}{n}2^{-2n} \\
&= \frac{2n}{2n - 1}\frac{(2n)!}{n!n!}2^{-2n} - \binom{2n}{n}2^{-2n} = \binom{2n}{n}2^{-2n}\left(\frac{2n}{2n - 1} - 1\right) \\
&= \frac{1}{2n - 1}\mathbf{P}(S_{2n} = 0).
\end{aligned}$$

2nd answer. Direct computation:

$$\begin{aligned}
\mathbf{P}(\tau_0 = 2n) &= \mathbf{P}(S_{2n} = 0, S_1 \ldots S_{2n-1} \neq 0) = 2\mathbf{P}(S_{2n} = 0, S_1 \ldots S_{2n-1} \neq 0, S_{2n-1} = 1) \\
&= 2\mathbf{P}(X_{2n} = -1, S_1 \ldots S_{2n-1} \neq 0, S_{2n-1} = 1) \\
&= 2\mathbf{P}(X_{2n} = -1)\,\mathbf{P}(S_1 \ldots S_{2n-1} \neq 0, S_{2n-1} = 1) \\
&= \frac{1}{2n-1}\mathbf{P}(S_{2n-1} = 1) = \frac{2}{2n-1} N_{2n-1}(0,1)\, 2^{-2n} = \frac{1}{2n-1}\mathbf{P}(S_{2n} = 0).
\end{aligned}$$

3. Let S_n be symmetric simple r.w. ($p = q = 1/2$), and $S_0 = 0$, i.e.,

$$S_n = X_1 + \ldots + X_n, n \geq 1,$$

where X_i are independent identically distributed, $\mathbf{P}(X_i = 1) = \mathbf{P}(X_i = -1) = 1/2$.

a) Show that $\bar{S}_n = -S_n, n \geq 0$, is symmetric r.w. as well, that is the sequences $\{S_n, n \geq 0\}$, and $\{-S_n, n \geq 0\}$ are identically distributed. Hint: X_i and $-X_i$ have identical mass functions, and $-X_i$ are independent.

Answer. First note that X_i and $-X_i$ are identically distributed (they have the same mass function):

$$\begin{aligned}
\mathbf{P}(X_i = 1) &= \mathbf{P}(-X_i = 1) = 1/2, \\
\mathbf{P}(X_i = -1) &= \mathbf{P}(-X_i = -1) = 1/2.
\end{aligned}$$

Since X_k are independent, $-X_k$ are independent as well, and for any n, the mass functions of $(X_1, \ldots, X_n)$ and $(-X_1, \ldots, -X_n)$ are the same. Thus for any n any $h(X_1, \ldots, X_n)$ and $h(-X_1, \ldots, -X_n)$ are identically distributed (they have the same mass function). Therefore $S_n, n \geq 1$, and $-S_n, n \geq 1$, are identically distributed. By definition $-S_n = (-X_1) + \ldots + (-X_n), n \geq 1$, is a r.w. ($-S_0 = 0$).

b) For $b \neq 0$, set $\tau_b = \tau_b(S) = \min\{n > 0 : S_n = b\}$. Show that

$$\mathbf{P}(\tau_b < \tau_{-b}) = \mathbf{P}(\tau_{-b} < \tau_b) = 1/2.$$

Hint. Recall for any $a \neq 0$, $\mathbf{P}(\tau_a < \infty) = 1$. Since the sequences $\{S_n, n \geq 0\}$, and $\{-S_n, n \geq 0\}$ are identically distributed,

$$\mathbf{P}(\tau_b(S) < \tau_{-b}(S)) = \mathbf{P}(\tau_b(-S) < \tau_{-b}(-S)),$$

where

$$\begin{aligned}
\tau_b(S) &= \min\{n > 0 : S_n = b\}, \tau_b(-S) = \min\{n > 0 : -S_n = b\}, \\
\tau_{-b}(S) &= \min\{n > 0 : S_n = -b\}, \tau_{-b}(-S) = \min\{n > 0 : -S_n = -b\}.
\end{aligned}$$

Answer. Recall for any $a \neq 0$, $\mathbf{P}(\tau_a < \infty) = 1$. Since the sequences $\{S_n, n \geq 0\}$, and $\{-S_n, n \geq 0\}$ are identically distributed,

$$\mathbf{P}(\tau_b(S) < \tau_{-b}(S)) = \mathbf{P}(\tau_b(-S) < \tau_{-b}(-S)),$$

where

$$\begin{aligned}
\tau_b(S) &= \min\{n > 0 : S_n = b\}, \tau_b(-S) = \min\{n > 0 : -S_n = b\} = \tau_{-b}(S), \\
\tau_{-b}(S) &= \min\{n > 0 : S_n = -b\}, \tau_{-b}(-S) = \min\{n > 0 : -S_n = -b\} = \tau_b(S).
\end{aligned}$$

So,

$$\mathbf{P}\left(\tau_b(S) < \tau_{-b}(S)\right) = \mathbf{P}\left(\tau_{-b}(S) < \tau_b(S)\right).$$

Since $\Omega = \{\tau_b(S) < \tau_{-b}(S)\} \cup \{\tau_{-b}(S) < \tau_b(S)\}$, it follows that

$$\mathbf{P}\left(\tau_b(S) < \tau_{-b}(S)\right) = \mathbf{P}\left(\tau_{-b}(S) < \tau_b(S)\right) = 1/2.$$

c) Let $\sigma_k = \min\{n > 0 : S_n \notin (-k, k)\}$. Find $\mathbf{E}\left(S_{\sigma_k}\right)$ and $\mathrm{var}\left(S_{\sigma_k}\right)$. Hint: $\sigma_k = \min\{\tau_k, \tau_{-k}\}$ and $\mathbf{P}(\tau_k < \infty) = \mathbf{P}(\tau_{-k} < \infty) = 1$. What values S_{σ_k} and $S_{\sigma_k}^2$ take?

Answer. Since $\mathbf{P}(\tau_k < \infty) = \mathbf{P}(\tau_{-k} < \infty) = 1$, and $\sigma_k = \min\{\tau_k, \tau_{-k}\}$, it follows that $\mathbf{P}(\sigma_k < \infty) = 1$, and

$$S_{\sigma_k} = \begin{cases} k & \text{if } \tau_k < \tau_{-k} \\ -k & \text{if } \tau_{-k} < \tau_k. \end{cases}$$

Hence $\mathbf{E}\left(S_{\sigma_k}\right) = k(1/2) + (-k)(1/2) = 0$, and $\mathrm{var}\left(S_{\sigma_k}\right) = \mathbf{E}\left(S_{\sigma_k}^2\right) = k^2$.

4. a) Let $S_0 = a > 0$, $p = q = 1/2$. Let $\tau_0 = \min\{k \geq 0 : S_k = 0\}$, the hitting time of zero. For $a \geq 1, j \geq 1, n \geq 1$, express

$$\mathbf{P}\left(S_n = j, \tau_0 \leq n | S_0 = a\right) \text{ and } \mathbf{P}\left(S_n = j, \tau_0 > n | S_0 = a\right)$$

in terms of finitely many basic probabilities. By basic probabilities for S_n,$n \geq 0$, we mean probabilities of the form

$$\mathbf{P}(S_n = k | S_0 = 0), \mathbf{P}(S_n \geq k | S_0 = 0), \mathbf{P}(S_n \leq k | S_0 = 0)$$

Hint. Reflection principle, and space homogeneity.

Answer. The event $\{S_n = j, \tau_0 \leq n\}$ consists of all paths from $(0, a)$ to (n, j) that cross or touch zero (x-axis) at some time moment $k \in \{0, 1, \ldots, n\}$. Let $N_n^0(a, j)$ be the number of such paths. Then $\mathbf{P}(S_n = j, \tau_0 \leq n | S_0 = a) = N_n^0(a, j) 2^{-n}$.By reflection principle, $N_n^0(a, j) = N_n(-a, j)$, and by space homogeneity,

$$\begin{aligned} \mathbf{P}\left(S_n = j, \tau_0 \leq n | S_0 = a\right) &= N_n^0(a, j) 2^{-n} = N_n(-a, j) 2^{-n} \\ &= \mathbf{P}\left(S_n = j | S_0 = -a\right) = \mathbf{P}\left(S_n = j + a | S_0 = 0\right). \end{aligned}$$

Then

$$\begin{aligned} & \mathbf{P}\left(S_n = j, \tau_0 > n | S_0 = a\right) \\ = & \mathbf{P}\left(S_n = j | S_0 = a\right) - \mathbf{P}\left(S_n = j, \tau_0 \leq n | S_0 = a\right) \\ = & \mathbf{P}\left(S_n = j - a | S_0 = 0\right) - \mathbf{P}\left(S_n = j + a | S_0 = 0\right). \end{aligned}$$

b) For $a \geq 1, j \geq 1, n \geq 1$, show that

$$\mathbf{P}\left(\tau_0 > n | S_0 = a\right) = \sum_{j=1-a}^{a} \mathbf{P}\left(S_n = j | S_0 = 0\right).$$

Hint. Remove S_n in joint probabilities of a); telescopic sums?

Answer. Since S_n takes only positive values if $\{\tau_0 > n\}$, changing the variable of summation,

$$\begin{aligned}
& \mathbf{P}\left(\tau_0 > n | S_0 = a\right) \\
= & \sum_{j=1}^{\infty} \mathbf{P}\left(S_n = j, \tau_0 > n | S_0 = a\right) = \sum_{j=1}^{\infty} \mathbf{P}\left(S_n = j - a | S_0 = 0\right) - \sum_{j=1}^{\infty} \mathbf{P}\left(S_n = j + a | S_0 = 0\right) \\
= & \sum_{k=1-a}^{\infty} \mathbf{P}\left(S_n = k | S_0 = 0\right) - \sum_{k=a+1}^{\infty} \mathbf{P}\left(S_n = k | S_0 = 0\right) = \sum_{k=1-a}^{a} \mathbf{P}\left(S_n = k | S_0 = 0\right).
\end{aligned}$$

5. In an election, candidate A receives n votes and candidate B receives m votes, where $n > m$. Assuming that all $\binom{m+n}{m} = \binom{m+n}{\frac{m+n+(n-m)-0}{2}}$ orderings (orders in which those votes were cast) are equally likely, show that probability that A is always ahead in the count of the votes is $\frac{n-m}{n+m}$.

Hint. $m+n$ votes were cast one by one. Let $\mathcal{H}_k$ be number of votes for A and $\mathcal{T}_k$ be number of votes for B after k people voted. Think about the differences $S_k = \mathcal{H}_k - \mathcal{T}_k, k \geq 1, S_0 = 0$.

Answer. Let $\mathcal{H}_k$ be number of votes for A and $\mathcal{T}_k$ be number of votes for B after k people voted. Consider the differences $S_k = \mathcal{H}_k - \mathcal{T}_k, 1 \leq k \leq m+n, S_0 = 0$. The "paths" of S_k are like the paths of the simple random walk. We know that the final difference $S_{n+m} = n - m$, and assume that all orders in which the votes were cast are equally likely. The number of those orders coincides with number of paths of S that start at 0 and reach $n-m$ in $n+m$ steps. That total number is

$$N_{m+n}\left(0, n-m\right) = \binom{n+m}{\frac{(n+m)+(n-m)-0}{2}} = \binom{n+m}{n} = \binom{n+m}{m},$$

that is there are $\binom{n+m}{m}$ equally likely paths (voting orders). The number of paths in the event $\{\mathcal{H}_1 > \mathcal{T}_1, \ldots, \mathcal{H}_{n+m} > \mathcal{T}_{n+m}, S_{n+m} = n-m\}$ is

$$\begin{aligned}
& \#\left\{\mathcal{H}_1 > \mathcal{T}_1, \ldots, \mathcal{H}_{n+m} > \mathcal{T}_{n+m}, S_{n+m} = n-m\right\} \\
= & \#\left\{S_1 \ldots S_{n+m} \neq 0, S_{n+m} = n-m\right\} = \overline{N_{n+m}^0}\left(0, n-m\right) =
\end{aligned}$$

= # of paths that connect in in $n+m$ steps zero and $n-m$ without revisiting zero. By Theorem we considered (see 3.10),

$$\overline{N_{n+m}^0}\left(0, n-m\right) = \frac{n-m}{n+m} N_{m+n}\left(0, n-m\right).$$

Thus

$$\begin{aligned}
& \mathbf{P}\left(\text{A is always ahead in the count of votes}\right) \\
= & \frac{\overline{N_{n+m}^0}\left(0, n-m\right)}{\binom{n+m}{m}} = \frac{\frac{n-m}{n+m} N_{m+n}\left(0, n-m\right)}{N_{m+n}\left(0, n-m\right)} = \frac{n-m}{n+m}.
\end{aligned}$$

2nd answer. The total number of equally likely orders is $\binom{m+n}{m}$. Then, by reflection principle,

$$\begin{aligned}
\overline{N^0_{n+m}}(0, n-m) &= \overline{N^0_{n+m-1}}(1, n-m) = N_{n+m-1}(1, n-m) - N^0_{n+m-1}(1, n-m) \\
&= N_{n+m-1}(1, n-m) - N_{n+m-1}(-1, n-m) \\
&= \binom{n+m-1}{\frac{(n+m-1)+(n-m)-1}{2}} - \binom{n+m-1}{\frac{(n+m-1)+(n-m)+1}{2}} = \binom{n+m-1}{n-1} - \binom{n+m-1}{n} \\
&= \frac{(n+m-1)!}{(n-1)!m!} - \frac{(n+m-1)!}{n!\,(m-1)!} = \frac{n-m}{n+m}\frac{(n+m)!}{n!m!} = \frac{n-m}{n+m}\binom{m+n}{m},
\end{aligned}$$

and

$$\begin{aligned}
&\mathbf{P}\,(\text{A is always ahead in the count of votes}) \\
&= \frac{\frac{n-m}{n+m}\binom{m+n}{m}}{\binom{m+n}{m}} = \frac{n-m}{n+m}.
\end{aligned}$$

9 hw9

1. Let S_n be a simple r.w. Assume $S_0 = 0$. Let $\tau_0 = \min\{k > 0 : S_k = 0\}$, the first return time back to zero. Show that

$$1 = \sum_{j=0}^{n} \mathbf{P}(\tau_0 > j)\, \mathbf{P}\left(S_{n-j} = 0\right).$$

Hint. $\mathbf{P}(\sigma_n = j)$, where $\sigma_n = \max\{k \le n : S_k = 0\}$?

Answer. We showed in class (posted note of 10/24) that

$$\mathbf{P}(\sigma_n = j) = \mathbf{P}(\tau_0 > n - j)\, \mathbf{P}\left(S_j = 0\right),$$

$j = 0, \ldots, n$: for $j < n$, using space homogeneity,

$$\begin{aligned}
\mathbf{P}(\sigma_n = j) &= \mathbf{P}\left(S_{j+1} \ldots S_n \neq 0, S_j = 0\right) = \mathbf{P}\left(S_{j+1} \ldots S_n \neq 0 | S_j = 0\right) \mathbf{P}\left(S_j = 0\right) \\
&= \mathbf{P}\left(S_1 \ldots S_{n-j} \neq 0 | S_0 = 0\right) \mathbf{P}\left(S_j = 0\right) = \mathbf{P}(\tau_0 > n - j)\, \mathbf{P}\left(S_j = 0\right).
\end{aligned}$$

Since by definition σ_n takes in $\{0, 1, \ldots, n\}$, we obtain (changing the variable of summation as well),

$$\begin{aligned}
1 &= \sum_{j=0}^{n} \mathbf{P}(\sigma_n = j) = \sum_{j=0}^{n} \mathbf{P}(\tau_0 > n - j)\, \mathbf{P}\left(S_j = 0\right) \\
&= \sum_{i=0}^{n} \mathbf{P}(\tau_0 > i)\, \mathbf{P}(S_{n-i} = 0).
\end{aligned}$$

2. The annual rainfall figures in Bandrika are independent identically distributed continuous random variables $\{X_r, r \ge 1\}$. Find the probability that:

(a) $X_1 < X_2 < X_3 < X_4$.

Answer. Since X_r are continuous,

$$\Omega = \cup_{i_1, i_2, i_3, i_4} \{X_{i_1} < X_{i_2} < X_{i_3} < X_{i_4}\}$$

The union of sets is taken over all different orderings of the numbers $\{1, 2, 3, 4\}$: there are 4! of them. Since X_1, X_2, X_3, X_4 are identically distributed and independent any ordering is equally likely (write corresponding quadruple integral for the probability and change the variables of integration $y_1 = x_{i_1}, \ldots, y_4 = x_{i_4}$):

$$4!\mathbf{P}(X_{i_1} < X_{i_2} < X_{i_3} < X_{i_4}) = 1, \mathbf{P}(X_{i_1} < X_{i_2} < X_{i_3} < X_{i_4}) = \frac{1}{4!}.$$

(b) $X_1 > X_2 < X_3 < X_4$. Hint. One way is to enumerate all possibilities and rewrite this event as a disjoint union of (a) type. The other way (without thinking): with probability 1,

$$\begin{aligned}
I_{\{X_1 > X_2 < X_3 < X_4\}} &= I_{\{X_1 > X_2\}} I_{\{X_2 < X_3\}} I_{\{X_3 < X_4\}} \\
&= I_{\{X_1 > X_2\}} \left(1 - I_{\{X_2 > X_3\}}\right) \left(1 - I_{\{X_3 > X_4\}}\right).
\end{aligned}$$

Answer. 1st answer:

$$\begin{aligned}
\mathbf{P}(X_1 > X_2 < X_3 < X_4) &= \mathbf{P}(X_2 < X_3 < X_4) - \mathbf{P}(X_1 < X_2 < X_3 < X_4) \\
&= \frac{1}{3!} - \frac{1}{4!} = \frac{3}{4!}.
\end{aligned}$$

2nd answer. We represent as disjoint union:

$$\begin{aligned} & \{X_1 > X_2 < X_3 < X_4\} \\ = \; & \{X_2 < X_1 < X_3 < X_4\} \cup \{X_2 < X_3 < X_1 < X_4\} \cup \{X_2 < X_3 < X_4 < X_1\}. \end{aligned}$$

By part a),

$$\mathbf{P}(X_1 > X_2 < X_3 < X_4) = 3 \cdot \frac{1}{4!} = \frac{3}{4!}.$$

3rd answer:

$$\begin{aligned} I_{\{X_1 > X_2 < X_3 < X_4\}} &= I_{\{X_1 > X_2\}} I_{\{X_2 < X_3 < X_4\}} = \left(1 - I_{\{X_1 < X_2\}}\right) I_{\{X_2 < X_3 < X_4\}} \\ &= I_{\{X_2 < X_3 < X_4\}} - I_{\{X_1 < X_2\}} I_{\{X_2 < X_3 < X_4\}} = I_{\{X_2 < X_3 < X_4\}} - I_{\{X_1 < X_2 < X_3 < X_4\}}, \end{aligned}$$

and we arrive at the first answer:

$$\begin{aligned} \mathbf{P}(X_1 > X_2 < X_3 < X_4) &= \mathbf{E} I_{\{X_2 < X_3 < X_4\}} - \mathbf{E} I_{\{X_1 < X_2 < X_3 < X_4\}} \\ &= \mathbf{P}(X_2 < X_3 < X_4) - \mathbf{P}(X_1 < X_2 < X_3 < X_4) \\ &= \frac{1}{3!} - \frac{1}{4!} = \frac{3}{4!} \end{aligned}$$

3. a) Let Θ be uniform on $(0, 2\pi)$, and $a > 0$. Find the pdf of $Y = a \cos \Theta$. Hint. Find pdf of $X = \cos \Theta$ first.

Answer. First for $z \in [-1, 1]$, looking at the graph of $\cos \theta, 0 \le \theta \le 2\pi$, we see that

$$\{\cos \Theta \le z\} = \left\{\cos^{-1} z \le \Theta \le 2\pi - \cos^{-1} z\right\}.$$

Since Θ is uniform in $(0, 2\pi)$,

$$G(z) = \mathbf{P}(\cos \Theta \le z) = \frac{2\pi - \cos^{-1} z - \cos^{-1} z}{2\pi} = 1 - \frac{\cos^{-1} z}{\pi}.$$

Hence the pdf

$$G'(z) = \frac{d}{dz}\left(\mathbf{P}(\cos \Theta \le z)\right) = \frac{1}{\pi} \frac{1}{\sqrt{1 - z^2}}, z \in (-1, 1).$$

Now, Y takes values in $[-|a|, |a|]$. For $y \in [-|a|, |a|]$,

$$\mathbf{P}(Y \le y) = \begin{cases} \mathbf{P}(\cos \Theta \le y/|a|) = G(y/|a|) & \text{if } a > 0, \\ \mathbf{P}(\cos \Theta \ge -y/|a|) = 1 - G(-y/|a|) & \text{if } a < 0. \end{cases}$$

Either way, the pdf, by differentiating,

$$f_Y(y) = \frac{1}{|a|} G'(y/|a|) = \frac{1}{\pi |a|} \frac{1}{\sqrt{1 - (y/|a|)^2}} = \frac{1}{\pi} \frac{1}{\sqrt{a^2 - y^2}}, y \in [-|a|, |a|].$$

b) Let X be uniform in $(0, 1)$. Find the df and pdf of $U = 1 - X$.

Answer. The range of U is $(0, 1)$. For $u \in (0, 1)$, the df

$$F(u) = \mathbf{P}(U \le u) = \mathbf{P}(1 - X \le u) = \mathbf{P}(X \ge 1 - u) = \frac{1 - (1 - u)}{1} = u,$$

$\mathbf{F}(u) = 0$ if $u \leq 0$, and $F(u) = 1$ if $u \geq 1$. Thus the pdf $f(u) = F'(u) = 1$ if $u \in (0,1)$ (zero otherwise). Thus $U = 1 - X$ is uniform in $(0,1)$.

c) Let X, Y be independent uniform in $(0,1)$. Find the df and pdf of $V = X - Y$. Hint. Determine the areas under $x - y \leq v$.

Answer. The range of $V = X - Y$ is $(-1,1)$. Drawing a picture and finding some triangle areas, we see that for $v \in -(1,0]$, the df

$$F(v) = \mathbf{P}(V \leq v) = \mathbf{P}(X \leq v + Y) = \frac{1}{2}(1+v)^2,$$

For $v \in (0,1)$,

$$F(v) = 1 - \frac{1}{2}(1-v)^2,$$

$F(v) = 0$ if $v \leq -1$ and $F(v) = 1$ if $v \geq 1$. The pdf $f(v) = 1 + v$ if $v \in (-1,0)$, $f(v) = 1 - v$ if $v \in (0,1)$ (zero otherwise).

4. *Order statistics*. Let $X_1, \ldots, X_n$ be independent identically distributed variables with a common pdf f. Such a collection is called a random sample. For each $\omega \in \Omega$, arrange the sample values $X_1(\omega), \ldots, X_n(\omega)$ in non-decreasing order $X_{(1)}(\omega), \ldots, X_{(n)}(\omega)$, where (1), (2),..., (n) is a (random) permutation (arrangement) of 1, 2,..., n. The new variables $X_{(1)}, \ldots, X_{(n)}$ are called the order statistics.

a) Show, by a symmetry argument, that the joint distribution function of the order statistics satisfies

$$\begin{aligned}
& \mathbf{P}\left(X_{(1)} \leq y_1, \ldots, X_{(n)} \leq y_n\right) \\
= \ & n!\mathbf{P}(X_1 \leq y_1, \ldots, X_n \leq y_n, X_1 < \ldots < X_n) \\
= \ & n!\int_{-\infty}^{y_1} \ldots \int_{-\infty}^{y_n} \chi(x_1, \ldots, x_n) f(x_1) \ldots f(x_n)\, dx_1 \ldots dx_n,
\end{aligned}$$

where

$$\chi(x_1, \ldots, x_n) = \begin{cases} 1 & \text{if } x_1 < \ldots < x_n \\ 0 & \text{otherwise.} \end{cases}$$

Note $X_{(1)} = \min_{1 \leq k \leq n} X_k$, $X_{(n)} = \max_{1 \leq k \leq n} X_k$.

Hint. We have

$$\left\{X_{(1)} \leq y_1, \ldots, X_{(n)} \leq y_n\right\} = \cup_{j_1,\ldots,j_n}\left\{X_{j_1} \leq y_1, \ldots, X_{j_n} \leq y_n, X_{j_1} < \ldots < X_{j_n}\right\},$$

where the union is taken over all possible different orderings (permutations) $j, \ldots, j_n$ of $\{1, \ldots, n\}$. All the sets in the union are disjoint, and there are $n!$ of them. With probability 1,

$$\Omega = \cup_{j_1,\ldots,j_n}\left\{X_{j_1} < \ldots < X_{j_n}\right\}.$$

Answer. With probability 1,

$$\Omega = \cup_{j_1,\ldots,j_n}\left\{X_{j_1} < \ldots < X_{j_n}\right\},$$

where the union is taken over all possible different orderings (permutations) $j, \ldots, j_n$ of $\{1, \ldots, n\}$. All the sets in the union are disjoint, and there are $n!$ of them. Hence

$$\left\{X_{(1)} \leq y_1, \ldots, X_{(n)} \leq y_n\right\} = \cup_{j_1,\ldots,j_n}\left\{X_{j_1} \leq y_1, \ldots, X_{j_n} \leq y_n, X_{j_1} < \ldots < X_{j_n}\right\},$$

and, by symmetry (write a multiple integral for probability and change the variables of integration, $y_1 = x_{j_1}, \dots, y_n = x_{j_n}$, we see that

$$\mathbf{P}\left(X_{j_1} \le y_1, \dots, X_{j_n} \le y_n, X_{j_1} < \dots < X_{j_n}\right)$$

is the same for any ordering $j, \dots, j_n$. Hence

$$\begin{aligned} & \mathbf{P}\left(X_{(1)} \le y_1, \dots, X_{(n)} \le y_n\right) \\ = \;& n!\mathbf{P}\left(X_1 \le y_1, \dots, X_n \le y_n, X_1 < \dots < X_n\right) \\ = \;& n! \int_{-\infty}^{y_1} \dots \int_{-\infty}^{y_n} \chi\left(x_1, \dots, x_n\right) f\left(x_1\right) \dots f\left(x_n\right) dx_1 \dots dx_n. \end{aligned} \tag{9.5}$$

b) Find the marginal density function of the kth order statistic $X_{(k)}$ of a sample with size n directly (without using the joint df find in a)). Hint. First, show that the df of $X_{(k)}$ is

$$F_{X_{(k)}}(x) = \mathbf{P}\left(X_{(k)} \le x\right) = \sum_{j=k}^{n} \binom{n}{j} F(x)^j \left[1 - F(x)\right]^{n-j} :$$

note $\sum_{j=1}^{n} I_{\{X_j \le x\}}$ is a binomial r.v.

Answer. Since X_j are identically distributed, $\mathbf{P}\left(X_j \le x\right) = F(x) \in (0,1)$, and $Y = \sum_{j=1}^{n} I_{\{X_j \le x\}}$, as the number of successes in n independent trials, is binomial with parameters n and $p = F(x)$. Hence

$$F_{X_{(k)}}(x) = \mathbf{P}\left(X_{(k)} \le x\right) = \mathbf{P}(Y \ge k) = \sum_{j=k}^{n} \binom{n}{j} F(x)^j \left[1 - F(x)\right]^{n-j},$$

and the pdf of $X_{(k)}$ is for $k < n$,

$$f_{X_{(k)}}(x) = F'_{X_{(k)}}(x) = f(x) \left\{ \sum_{j=k}^{n} \binom{n}{j} jF(x)^{j-1} \left[1 - F(x)\right]^{n-j} - \binom{n}{j} (n-j) F(x)^j \left[1 - F(x)\right]^{n-j-1} \right\}.$$

Comment. This formula can be simplified. Since

$$\binom{n}{j} j = n \frac{(n-1)!}{(j-1)!\,(n-j)!} = n \binom{n-1}{j-1}, \binom{n}{j} (n-j) = n \frac{(n-1)!}{j!\,(n-j-1)!} = n \binom{n-1}{j}$$

with $\binom{n-1}{j} := 0$ if $j = n$, and

$$\sum_{j=k}^{n} \left(a_{j-1} - a_j\right) = a_{k-1} - a_n,$$

we have

$$\begin{aligned} f_{X_{(k)}}(x) &= nf(x) \left\{ \sum_{j=k}^{n} \binom{n-1}{j-1} F(x)^{j-1} \left[1 - F(x)\right]^{n-j} - \binom{n-1}{j} F(x)^j \left[1 - F(x)\right]^{n-j-1} \right\} \\ &= nf(x) \binom{n-1}{k-1} F(x)^{k-1} \left[1 - F(x)\right]^{n-k} = \binom{n}{k} kF(x)^{k-1} f(x) \left[1 - F(x)\right]^{n-k}. \end{aligned}$$

c) Find the joint density function of the order statistics of n independent uniform variables in $(0, T)$. Hint: use a)

Answer. According to part a) (look at (9.5)) and definition of a jointly continuous r. vector, the joint pdf for $Y = \left(X_{(1)}, \ldots, X_{(n)}\right)$ is

$$f_Y(y) = n!\chi(y_1, \ldots, y_n) f(y_1) \ldots f(y_n).$$

For X, uniform in $(0, T)$, we have $f(x) = 1/T = T^{-1}$ if $x \in (0, T)$ or $f(x) = T^{-1} I_{(0,T)}(x)$. Since in this case the product $f(y_1) \ldots f(y_n) \neq 0$ iff $y_1, \ldots, y_n \in (0, T)$, we have

$$\begin{aligned} f_Y(y) &= n!\chi(y_1, \ldots, y_n) T^{-n}, y_1, \ldots, y_n \in (0, T), \text{ equivalently,} \\ f_Y(y) &= n! T^{-n}, 0 < y_1 < \ldots < y_n < T, \end{aligned}$$

meaning

$$f_Y(y) = \begin{cases} \frac{n!}{T^n} & \text{if } 0 < y_1 < \ldots < y_n < T, \\ 0 & \text{otherwise.} \end{cases}$$

Hence Y **is uniform** in $D = \{(y_1, \ldots, y_n) : 0 < y_1 < \ldots < y_n < T\}$.

5. Let $X_1, \ldots, X_n$ be positive independent identically distributed continuous random variables. Show that, if $m < n$, then $\mathbf{E}\left(\frac{S_m}{S_n}\right) = \frac{m}{n}$, where $S_k = X_1 + \ldots + X_k$.

Hint. The random variables $\frac{X_i}{S_n} = \frac{X_i}{X_1 + \ldots + X_n}, i = 1 \ldots, n$, are identically distributed. First think what is $\mathbf{E}\left(\frac{X_i}{S_n}\right)$?

Answer. Let $f(x)$ be the pdf of X_1. Then the product $f(x_1) \ldots f(x_n)$ is the joint pdf of $(X_1, \ldots, X_n)$. Then for any $r \in \mathbf{R}$, the distribution function

$$\mathbf{P}\left(\frac{X_i}{S_n} \leq r\right) = \int_{\left\{\frac{x_i}{x_1 + \ldots + x_n} \leq r\right\}} \cdots \int f(x_1) \ldots f(x_n) \, dx_1 \ldots dx_n$$

is independent of i (look at the multiple integral). Hence (X_i / S_n) are identically distributed. Since $\sum_{i=1}^n \frac{X_i}{S_n} = 1$, we have immediately that

$$\mathbf{E}\left(\frac{X_i}{S_n}\right) = \frac{1}{n}, i = 1, \ldots, n.$$

Thus

$$\mathbf{E}\left(\frac{S_m}{S_n}\right) = \mathbf{E}\left(\sum_{i=1}^m \frac{X_i}{S_n}\right) = \sum_{i=1}^m \mathbf{E}\left(\frac{X_i}{S_n}\right) = \frac{m}{n}.$$

10 hw10

1. A coin-making machine produces quarters in such way that, for each coin, the probability U to turn up heads is uniform in $(0,1)$. A coin pops out (randomly) and is flipped 10 times. Let X be number of heads in those 10 tosses.

a) What are

$$\begin{aligned}\mathbf{P}(X=j|U=u), j &= 0,\dots,10 \text{ and}\\ \mathbf{P}(X=j|U), j &= 0,\dots,10?\end{aligned}$$

Hint. X ~binomial$(10, U)$.

Answer. Given $U=u$, the r.v. X ~bin$(10, U)$:

$$\mathbf{P}(X=j|U=u)=\mathbf{E}\left(I_{\{j\}}(X)|U=u\right)=\binom{10}{j}u^j(1-u)^{10-j}, j=0,\dots,10.$$

Hence

$$\mathbf{P}(X=j|U)=\mathbf{E}\left(I_{\{j\}}(X)|U\right)=\binom{10}{j}U^j(1-U)^{10-j}, j=0,\dots,10.$$

b) Find $\mathbf{P}(X=j), j=0,\dots,10$, the distribution of X. Hint: $\mathbf{P}(X=j)=\mathbf{E}[\mathbf{P}(X=j|U)]$, $j=0,\dots,10$; It is known that

$$\int_0^1 u^j(1-u)^{n-j}\,du=\frac{j!(n-j)!}{(n+1)!}.$$

Answer. By part a), for $j=0,1,\dots,10$,

$$\begin{aligned}\mathbf{P}(X=j) &= \mathbf{E}\left(I_{\{j\}}(X)\right)=\mathbf{E}\left[\mathbf{E}\left(I_{\{j\}}(X)|U\right)\right]=\mathbf{E}[\mathbf{P}(X=j|U)]\\ &= \binom{10}{j}\mathbf{E}\left[U^j(1-U)^{10-j}\right]=\frac{10!}{j!(10-j)!}\int_0^1 u^j(1-u)^{10-j}\,du\\ &= \frac{10!}{j!(10-j)!}\frac{j!(10-j)!}{11!}=\frac{1}{11}.\end{aligned}$$

That is X is uniformly distributed in $\{0,1,\dots,10\}$.

c) Find $\mathbf{E}(X)$ and var(X).

Answer. Given $U=u$, the r.v. X ~bin$(10,u)$. Since

$$\begin{aligned}\mathbf{E}(X|U=u) &= 10u,\ \operatorname{var}(X|U=u)=10u(1-u),\\ \mathbf{E}\left(X^2|U=u\right) &= 10u(1-u)+(10u)^2=10u+90u^2.\end{aligned}$$

Hence $\mathbf{E}(X|U)=10U, \mathbf{E}\left(X^2|U\right)=10U+90U^2$

$$\begin{aligned}\mathbf{E}(X) &= 10\mathbf{E}(U)=10/2=5,\\ \mathbf{E}\left(X^2\right) &= 10\mathbf{E}(U)+90\mathbf{E}\left(U^2\right)=5+90/3=35,\end{aligned}$$

because $\mathbf{E}\left(U^2\right)=\int_0^1 u^2du=\frac{1}{3}$. Thus Var$(X)=35-5^2=10$.

Alternatively, we could compute directly using the pmf of X.

2. a) Let X, Y be independent standard normal. What is joint pdf of (X, Y)? Are $R^2 = X^2 + Y^2$ and $V = \frac{X}{\sqrt{X^2+Y^2}}$ independent? Hint. Use polar coordinates to find the joint df of $\left(R^2, V\right)$.

Answer. Because of independence, the joint pdf

$$f(x, y) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} = \frac{1}{2\pi} \exp\left\{-\frac{1}{2}\left(x^2 + y^2\right)\right\}, x, y \in \mathbf{R}.$$

All the values of V are between -1 and 1 and the range of R is $(0, \infty)$. For $v \in [-1, 1], u > 0$,

$$\{R \le u, V \le v\} = \left\{X^2 + Y^2 \le u, \frac{X}{\sqrt{X^2 + Y^2}} \le v\right\}.$$

Hence

$$\mathbf{P}(R \le u, V \le v) = \int\int_D f(x, y)\, dxdy = \frac{1}{2\pi}\int\int_D \exp\left\{-\frac{1}{2}\left(x^2 + y^2\right)\right\} dxdy,$$

where

$$D = \left\{(x, y) : x^2 + y^2 \le u, \frac{x}{\sqrt{x^2 + y^2}} \le v\right\}.$$

Using polar coordinates $x = r\cos\theta, y = r\sin\theta$, for the double integral, for any $u > 0, v \in [-1, 1]$,

$$\begin{aligned}
&\mathbf{P}(R \le u, V \le v) \\
&= \frac{1}{2\pi}\int_0^{\sqrt{u}}\int_0^{2\pi} I_{\{\cos\theta \le v\}} \exp\left\{-\frac{1}{2}r^2\right\} r dr d\theta \\
&= \left(\int_0^{\sqrt{u}} \exp\left\{-\frac{1}{2}r^2\right\} r dr\right)\left(\frac{1}{2\pi}\int_0^{2\pi} I_{\{\cos\theta \le v\}} d\theta\right).
\end{aligned}$$

Note that

$$1 = \mathbf{P}(R < \infty, V \le 1) = \int_0^{\infty} \exp\left\{-\frac{1}{2}r^2\right\} r dr,$$

and for any $u > 0, v \in [-1, 1]$.

$$\begin{aligned}
&\mathbf{P}(R \le u) = \mathbf{P}(R \le u, V \le 1) \\
&= \left(\int_0^{\sqrt{u}} \exp\left\{-\frac{1}{2}r^2\right\} r dr\right)\left(\frac{1}{2\pi}\int_0^{2\pi} I_{\{\cos\theta \le 1\}} d\theta\right) = \int_0^{\sqrt{u}} \exp\left\{-\frac{1}{2}r^2\right\} r dr,
\end{aligned}$$

$$\begin{aligned}
\mathbf{P}(V \le v) &= \mathbf{P}(V \le v, R < \infty) = \left(\int_0^{\infty} \exp\left\{-\frac{1}{2}r^2\right\} r dr\right)\left(\frac{1}{2\pi}\int_0^{2\pi} I_{\{\cos\theta \le v\}} d\theta\right) \\
&= \frac{1}{2\pi}\int_0^{2\pi} I_{\{\cos\theta \le v\}} d\theta,
\end{aligned}$$

and

$$\mathbf{P}(R \le u, V \le v) = \mathbf{P}(R \le u)\,\mathbf{P}(V \le v),$$

i.e. R, V are independent.

b) Let X, Y, Z be independent standard normal. Show that

$$\frac{X+YZ}{\sqrt{1+Z^2}} \text{ and } Z$$

are independent standard normal. Hint: condition on Z.

Answer. Given $Z=z$,

$$\frac{X+YZ}{\sqrt{1+Z^2}} = \frac{X+Yz}{\sqrt{1+z^2}} = \frac{1}{\sqrt{1+z^2}}X + \frac{z}{\sqrt{1+z^2}}Y.$$

As a linear combination of independent normal r.v. it is normal with parameters

$$\begin{aligned} \mu &= \frac{1}{\sqrt{1+z^2}}\mathbf{E}(X) + \frac{z}{\sqrt{1+z^2}}\mathbf{E}(Y) = 0, \\ \sigma^2 &= \operatorname{var}\left(\frac{1}{\sqrt{1+z^2}}X + \frac{z}{\sqrt{1+z^2}}Y\right) \\ &= \frac{1}{1+z^2}(1) + \frac{z^2}{1+z^2}(1) = 1, \end{aligned}$$

i.e. given $Z=z$, $\frac{X+YZ}{\sqrt{1+Z^2}}$ is standard normal

$$\begin{aligned} \mathbf{P}\left(\frac{X+YZ}{\sqrt{1+Z^2}} \le u|Z\right) &= \mathbf{E}\left[I_{\left\{\frac{X+YZ}{\sqrt{1+Z^2}}\le u\right\}}|Z\right] = \Phi(u), \\ \mathbf{P}\left(\frac{X+YZ}{\sqrt{1+Z^2}} \le u\right) &= \mathbf{E}\left[I_{\left\{\frac{X+YZ}{\sqrt{1+Z^2}}\le u\right\}}\right] = \mathbf{E}\left[\mathbf{E}\left(I_{\left\{\frac{X+YZ}{\sqrt{1+Z^2}}\le u\right\}}|Z\right)\right] \\ &= \mathbf{E}\left[\mathbf{P}\left(\frac{X+YZ}{\sqrt{1+Z^2}} \le u|Z\right)\right] = \Phi(u). \end{aligned}$$

For any u, v,

$$\begin{aligned} \mathbf{P}\left(\frac{X+YZ}{\sqrt{1+Z^2}} \le u, Z \le v\right) &= \mathbf{E}\left[I_{\left\{\frac{X+YZ}{\sqrt{1+Z^2}}\le u\right\}} I_{\{Z\le v\}}\right] = \mathbf{E}\left[\mathbf{E}\left(I_{\left\{\frac{X+YZ}{\sqrt{1+Z^2}}\le u\right\}}|Z\right) I_{\{Z\le v\}}\right] \\ &= \Phi(u)\,\mathbf{E}\left(I_{\{Z\le v\}}\right) = \Phi(u)\,\Phi(v), \end{aligned}$$

that is $V = \frac{X+YZ}{\sqrt{1+Z^2}}$ and Z are independent standard normals.

3. Let (X_1, X_2) be normal bivariate with $\mu_1 = \mu_2 = 0, \sigma_1^2 = \sigma_2^2 = 1$, and $\rho = 3/5$. Let (Y_1, Y_2) be the midterm and final exam scores of a randomly selected student. Assume

$$Y_1 = 80 + 3X_1, Y_2 = 75 + 2X_2.$$

Given a student got 90 in the midterm exam,

(a) What is the expectation and variance of her final exam score? Hint. It might be easier to reduce the question to (X_1, X_2) but also (Y_1, Y_2) is a normal bivariate.

Answer. Since (X_1, X_2) is standard bivariate normal with $\rho = 3/5$, we see that given $Y_1 = 90$, equivalently, $X_1 = 10/3$, we have $X_2 \sim N\left(\frac{3}{5} \cdot \frac{10}{3}, \frac{16}{25}\right) = N\left(2, \frac{16}{25}\right)$. Hence

$$\mathbf{E}\left(Y_2 | X_1 = 10/3\right) = \mathbf{E}\left(75 + 2X_2 | X_1 = \frac{10}{3}\right) = 75 + 2 \cdot 2 = 79,$$

and $\mathrm{Var}\left(Y_2 | X_1 = \frac{10}{3}\right) = 2^2 \mathrm{Var}\left(X_2 | X_1 = \frac{10}{3}\right) = 4 \cdot \frac{16}{25} = \frac{64}{25}$.

(b) What is the probability that she got more than 75 in the final exam? Express the probability in terms of $\Phi(x)$, the df of a standard normal r.v.

Answer. Given $Y_2 = 90$, equivalently, given $X_1 = 10/3$, we have $X_2 \sim N\left(2, \frac{16}{25}\right)$. Hence

$$\begin{aligned}
\mathbf{P}\left(Y_2 > 75 | Y_1 = 90\right) &= \mathbf{P}\left(75 + 2X_2 > 75 | X_1 = \frac{10}{3}\right) \\
&= \mathbf{P}\left(X_2 > 0 | X_1 = \frac{10}{3}\right) = 1 - \Phi\left(\frac{0-2}{\sqrt{\frac{16}{25}}}\right) \\
&= 1 - \Phi\left(-\frac{5}{2}\right) = \Phi\left(\frac{5}{2}\right) \approx 0.99.
\end{aligned}$$

4. Let X, Y be normal bivariate r.v. with $\mu_1 = \mu_2 = 0$, variances σ_1^2, σ_2^2 and correlation coefficient ρ.

a) Write what are $\mathbf{E}(X|Y)$, $\mathrm{Var}(X|Y)$?

Answer. We found in class that

$$\begin{aligned}
\mathbf{E}(X|Y) &= \rho \frac{\sigma_1}{\sigma_2}\left(Y - \mu_2\right) + \mu_1 = \rho \frac{\sigma_1}{\sigma_2} Y \\
\mathrm{Var}(X|Y) &= \sigma_1^2\left(1 - \rho^2\right)
\end{aligned}$$

b) Show that

$$\begin{aligned}
\mathbf{E}(X | X + Y = z) &= \frac{\sigma_1^2 + \rho\sigma_1\sigma_2}{\sigma_2^2 + 2\rho\sigma_1\sigma_2 + \sigma_1^2} z, \\
\mathrm{Var}(X | X + Y = z) &= \frac{\sigma_1^2\sigma_2^2\left(1 - \rho^2\right)}{\sigma_2^2 + 2\rho\sigma_1\sigma_2 + \sigma_1^2}.
\end{aligned}$$

Hint. $(X, X + Y)$is normal bivariate: apply a).

Answer. Since $(X, X + Y) = (X, Y) A$, where

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \det(A) = 1 \neq 0,$$

the r. vector $\left(X, \tilde{Y}\right) = (X, X + Y)$ is normal bivariate. We find its parameters: $\mu_X = 0$, $\mathrm{Var}(X) = \sigma_1^2$, $\tilde{\mu} = \mathbf{E}\left(\tilde{Y}\right) = 0$; by variance-covariance expansion,

$$\begin{aligned}
\tilde{\sigma}_2^2 &= \mathrm{Var}\left(\tilde{Y}\right) = \sigma_1^2 + \sigma_2^2 + 2\mathrm{Cov}(X, Y) \\
&= \sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2,
\end{aligned}$$

and

$$\tilde{\rho} = \rho(X, X+Y) = \frac{\mathrm{Cov}(X, X+Y)}{\sigma_1 \tilde{\sigma}_2} = \frac{\sigma_1^2 + \rho\sigma_1\sigma_2}{\sigma_1\tilde{\sigma}_2} = \frac{\sigma_1 + \rho\sigma_2}{\tilde{\sigma}_2},$$

where $\tilde{\sigma}_2 = \sqrt{\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2}$.

Hence by a),

$$\mathbf{E}(X|X+Y=z) = \tilde{\rho}\frac{\sigma_1}{\tilde{\sigma}_2}z == \frac{(\sigma_1 + \rho\sigma_2)\sigma_1}{\tilde{\sigma}_2^2} = \frac{\sigma_1^2 + \rho\sigma_1\sigma_2}{\sigma_2^2 + 2\rho\sigma_1\sigma_2 + \sigma_1^2}z,$$

and

$$\mathrm{Var}(X|X+Y=z) = \sigma_1^2\left(1-\tilde{\rho}^2\right) = \frac{\sigma_1^2\sigma_2^2\left(1-\rho^2\right)}{\sigma_2^2 + 2\rho\sigma_1\sigma_2 + \sigma_1^2}$$

because

$$1-\tilde{\rho}^2 = 1 - \frac{(\sigma_1+\rho\sigma_2)^2}{\tilde{\sigma}_2^2} = \frac{\tilde{\sigma}_2^2 - (\sigma_1+\rho\sigma_2)^2}{\tilde{\sigma}_2^2} = \frac{\sigma_2^2\left(1-\rho^2\right)}{\sigma_1^2+\sigma_2^2+2\rho\sigma_1\sigma_2}.$$

5. Let $Y = X + \varepsilon Z$, where $X \sim N\left(0, \sigma_1^2\right), Z \sim N(0,1)$ are independent (X is a random "target", εZ is a "noise", Y is what we observe and register). Find the best mean square estimate of X based on Y (recall it is $\hat{X} = \mathbf{E}(X|Y)$). How is $X - \hat{X}$ distributed? Find the mean square error $\mathbf{E}\left[\left(X-\hat{X}\right)^2\right]$.

Answer. Assuming $\varepsilon \neq 0$, the r. vector $(X, U) = (X, \varepsilon Z)$ is normal bivariate with parameters $\mu_X = 0 = \mu_U, \sigma_X^2 = \sigma_1^2, \sigma_U^2 = \varepsilon^2, \tilde{\rho} = \rho(X, U) = 0$. Applying #4b) above,

$$\mathbf{E}(X|X+U=y) = \frac{\sigma_1^2}{\varepsilon^2+\sigma_1^2}y, \hat{X} = \mathbf{E}(X|Y) = \frac{\sigma_1^2}{\varepsilon^2+\sigma_1^2}Y.$$

Alternatively, $(X, Y) = (X, X+\varepsilon Z) = (X, Z)A$ with

$$A = \begin{pmatrix} 1 & 1 \\ 0 & \varepsilon \end{pmatrix}, \det(A) \neq 0,$$

that is (X, Y) is normal bivariate, and we could find its parameters and apply #4a).

The error, denoting $\sigma_2^2 = \mathrm{Var}(Y) = \sigma_1^2 + \varepsilon^2$,

$$X - \hat{X} = X - \frac{\sigma_1^2}{\sigma_2^2}(X+\varepsilon Z) = \left(1 - \frac{\sigma_1^2}{\sigma_2^2}\right)X - \frac{\sigma_1^2}{\sigma_2^2}\varepsilon Z,$$

and the mean square error is, by independence of X, Z,

$$\begin{aligned}\mathbf{E}\left[\left(X-\hat{X}\right)^2\right] &= \mathrm{Var}\left(X-\hat{X}\right) = \left(1-\frac{\sigma_1^2}{\sigma_2^2}\right)^2\sigma_1^2 + \left(\frac{\sigma_1^2}{\sigma_2^2}\right)^2\varepsilon^2 \\ &= \frac{\varepsilon^4\sigma_1^2}{\left(\sigma_1^2+\varepsilon^2\right)^2} + \frac{\sigma_1^4\varepsilon^2}{\left(\sigma_1^2+\varepsilon^2\right)^2} = \varepsilon^2\frac{\sigma_1^2}{\sigma_1^2+\varepsilon^2}.\end{aligned}$$

Comment. If Y itself is used, then the error

$$\mathbf{E}\left[(Y-X)^2\right] = \varepsilon^2 > \varepsilon^2\frac{\sigma_1^2}{\sigma_1^2+\varepsilon^2}.$$

11 hw11

1. Suppose (X, Y) has joint density of the form $f(x, y) = g\left(\sqrt{x^2 + y^2}\right)$, for some function g. Let $R = \sqrt{X^2 + Y^2}$, and Θ be the polar angle of (X, Y). Find joint pdf of (R, Θ). Are R, Θ independent? Identify the distribution of Θ.

Hint. Use polar coordinates in $\mathbf{R}^2 : x = r\cos\theta, y = r\sin\theta$ with $r > 0, \theta = \arctan\frac{y}{x} \in (0, 2\pi)$. Note $r^2 = x^2 + y^2$. Recall Calculus: if D is a polar rectangle,

$$D = \{(x, y) : a < r < b, \alpha < \theta < \beta\} \text{ with } 0 < a < b, 0 < \alpha < \beta < 2\pi,$$

then

$$\int\int_D h(x, y)\, dxdy = \int_a^b \int_\alpha^\beta h(r\cos\theta, r\sin\theta)\, rd\theta dr.$$

Hence if $f(x, y)$ is joint pdf of (X, Y), then with D above we have

$$\mathbf{P}(a < R < b, \alpha < \Theta < \beta) = \int\int_D f(x, y)\, dxdy.$$

Answer. According to the hint, using calculus, for any $0 < a < b, 0 < \alpha < \beta < 2\pi$,

$$\begin{aligned}
\mathbf{P}(a < R < b, \alpha < \Theta < \beta) &= \int\int_D f(x, y)\, dxdy \qquad (11.6)\\
&= \int_a^b \int_\alpha^\beta f(r\cos\theta, r\sin\theta)\, rd\theta dr = \int_a^b \int_\alpha^\beta g(r)\, rd\theta dr\\
&= \int_a^b g(r)\, rdr\ \ (\beta - \alpha)
\end{aligned}$$

According to the formula (11.6), the joint pdf of (R, Θ) is

$$h(r, \theta) = g(r)\, r, 0 < \theta < 2\pi, r > 0.$$

Then the marginal pdf

$$\begin{aligned}
f_R(r) &= \int_0^{2\pi} h(r, \theta)\, d\theta = \int_0^{2\pi} g(r)\, rd\theta = 2\pi g(r)\, r, r > 0,\\
h(r, \theta) &= 2\pi g(r)\, r \cdot \frac{1}{2\pi}, 0 < \theta < 2\pi, r > 0,\\
f_\Theta(\theta) &= \int_0^\infty h(r, \theta)\, dr = \frac{1}{2\pi}, 0 < \theta < 2\pi.
\end{aligned}$$

Hence Θ is uniform in $(0, 2\pi)$. Since $h(r, \theta) = f_R(r)\, f_\Theta(\theta)$, R, Θ are independent.

2. Let X_1, X_2, X_3 be independent exponential r.v. with parameter $\lambda = 1 : \mathbf{P}(X_i > x) = e^{-x}, x > 0, i = 1, 2, 3$. Let

$$Y_1 = \frac{X_1}{X_1 + X_2 + X_3}, Y_2 = \frac{X_1 + X_2}{X_1 + X_2 + X_3}, Y_3 = X_1 + X_2 + X_3.$$

a) Find the joint pdf of (Y_1, Y_2, Y_3) and $\mathbf{E}(Y_i), i = 1, 2, 3$. Hint: Find the inverse and Jacobian etc.

Answer. Range of $X = (X_1, X_2, X_3) = D = \{(x_1, x_2, x_3) : x_1 > 0, x_2 > 0, x_3 > 0\}$, and the joint pdf by independence is

$$f(x) = f(x_1, x_2, x_3) = e^{-x_1}e^{-x_1}e^{-x_1} = e^{-(x_1+x_2+x_3)}, x = (x_1, x_2, x_3) \in D.$$

Now, $Y = (Y_1, Y_2, Y_3) = H(X)$, where $y = H(x)$ is given by

$$\begin{aligned} y_1 &= \frac{x_1}{x_1 + x_2 + x_3}, y_2 = \frac{x_1 + x_2}{x_1 + x_2 + x_3}, y_3 = x_1 + x_2 + x_3, \\ (x_1, x_2, x_3) &\in D. \end{aligned} \tag{11.7}$$

Given $y = (y_1, y_2, y_3) \in S$, we find the inverse $G(y) = H^{-1}(y)$ by solving for $x = (x_1, x_2, x_3)$ the equation $y = H(x)$, that is (11.7) above. We find

$$\begin{aligned} x_1 &= y_1 y_3, \\ x_2 &= y_3 y_2 - x_1 = y_3 y_2 - y_1 y_3, \\ x_3 &= y_3 - (x_1 + x_2) = y_3 - y_2 y_3, \end{aligned} \tag{11.8}$$

that is $G(y) = H^{-1}(y) = (y_1 y_3, y_3 y_2 - y_1 y_3, y_3 - y_2 y_3), y \in S$, where S is found by using (11.8) to "translate" D into S :

$$\begin{aligned} S &= \{(y_1, y_2, y_3) : x_1 = y_1 y_3 > 0, x_2 = y_3(y_2 - y_1) > 0, x_3 = y_3(1 - y_2) > 0\} \\ &= \{(y_1, y_2, y_3) : 0 < y_1 < y_2 < 1, y_3 > 0\}. \end{aligned}$$

The Jacobian of the inverse with $y = (y_1, y_2, y_3), y_3 > 0, 0 < y_1 < y_2 < 1$,

$$J(y) = \begin{vmatrix} y_3 & 0 & y_1 \\ -y_3 & y_3 & y_2 - y_1 \\ 0 & -y_3 & 1 - y_2 \end{vmatrix} = y_3^2.$$

Hence the joint pdf of Y is

$$\begin{aligned} f_Y(y) &= f_X(G(y))\,|J(y)|\, I_S(y) = e^{-y_3} y_3^2 I_{\{0<y_1<y_2<1, y_3>0\}} \\ &= \left(\frac{e^{-y_3} y_3^2}{2!} I_{\{y_3>0\}}\right) 2! I_{\{0<y_1<y_2<1\}}. \end{aligned}$$

b) Show that (i) (Y_1, Y_2) is distributed as the order statistics of two uniform r.v in $(0, 1)$, , see #5c) of hw8; Y_3 is gamma$(1, 3)$-distributed; (ii) Y_3 and (Y_1, Y_2) are independent.

Answer. (i)-(ii) We know that Y_3 is $\Gamma(1, 3)$ as the sum of 3 independent exponential with parameter 1. Thus

$$f_{Y_3}(y_3) = \frac{e^{-y_3} y_3^2}{2!} I_{\{y_3>0\}}, f_Y(y) = f_{Y_3}(y_3)\, 2! I_{\{0<y_1<y_2<1\}}, y = (y_1, y_2, y_3) \in \mathbf{R}^3.$$

Hence

$$f_{Y_1,Y_2}(y_1, y_2) = \int_{-\infty}^{\infty} f_Y(y_1, y_2, y_3)\, dy_3 = 2! I_{\{0<y_1<y_2<1\}},$$

and

$$f_Y(y) = f_{Y_3}(y_3) f_{Y_1,Y_2}(y_1, y_2), y = (y_1, y_2, y_3) \in \mathbf{R}^3.$$

So, Y_3 and (Y_1, Y_2) are independent. By #4c of hw9, (Y_1, Y_2) is order statistics of two independent uniform r.v. in $(0, 1)$.

Alternatively, knowing that Y_3 is $\Gamma(1, 3)$ and noticing by #4c that $2! I_{\{0<y_1<y_2<1\}}$ is the pdf of order statistics of two independent uniform r.v. in $(0, 1)$, we see that

$$f_Y(y) = f_{Y_3}(y_3) f_{Y_1,Y_2}(y_1, y_2).$$

3. a) A $d \times d$ symmetric matrix $A = \left(a_{ij}\right)$ is called nonnegative if for any $z = (z_1, \ldots, z_d) \in \mathbf{R}^d$,

$$zAz' = \sum_{i,j=1}^{d} a_{ij} z_i z_j = \sum_{i=1}^{d} a_{ii} z_i^2 + 2 \sum_{i<j} a_{ij} z_i z_j \geq 0.$$

where $z = (z_1, \ldots, z_d)$ is a row vector, z' is the transpose of z.

Let $X = (X_1, \ldots, X_d)$ and $\mathbf{E}\left(X_i^2\right) < \infty, i = 1, \ldots, d$. Let $B = \left(b_{ij}\right)$ with $b_{ij} = \mathrm{Cov}\left(X_i, X_j\right), i, j = 1, \ldots, d$. Show that B is nonnegative definite.

Answer. For any $z = (z_1, \ldots, z_d)$, we have, by recognizing variance-covariance expansion,

$$\begin{aligned} zBz' &= \sum_{i,j=1}^{d} \mathrm{Cov}\left(X_i, X_j\right) z_i z_j = \sum_{i=1}^{d} \mathrm{Var}\left(X_i\right) z_i^2 + 2 \sum_{i<j} z_i z_j \mathrm{Cov}\left(X_i, X_j\right) \\ &= \mathrm{Var}\left(\sum_{i=1}^{d} z_i X_i\right) \geq 0. \end{aligned}$$

b) Let $X = (X_1, \ldots, X_d)$ be multivariate normal $N(\mu, B)$, where μ is the vector of expected values and $B = \left(b_{ij}\right)$ is the covariance matrix: $b_{ij} = \mathrm{Cov}\left(X_i, X_j\right)$.

Let $c_1, \ldots, c_d$ be constants, and at least one of them $\neq 0$. What is the distribution of $Y = c_1 X_1 + \ldots + c_d X_d$? Determine the parameters of that distribution.

Answer. Since X is multivariate normal and $c = (c_1, \ldots, c_2) \neq 0$, Y is normally distributed. We find

$$\begin{aligned} \mathbf{E}(Y) &= c_1 \mu_1 + \ldots + c_d \mu_d = c\mu', \\ \mathrm{Var}(Y) &= \sum_{i=1}^{d} c_i^2 \mathrm{Var}\left(X_i\right) + 2 \sum_{i<j} \mathrm{Cov}\left(X_i, X_j\right) c_i c_j \\ &= \sum_{i=1}^{d} b_{ii} c_i^2 + 2 \sum_{i<j} b_{ij} c_i c_j = cBc' > 0. \end{aligned}$$

So, $Y \sim N(c\mu', cBc')$.

4. a) Let X be uniform in $(0, 1)$. Find the df and pdf of $V = \min\{1 - X, X\}$.

Answer. Range of $V = (0, 1/2)$.

1st answer. For $v \in (0, 1/2)$,

$$\begin{aligned}\mathbf{P}(V \le v) &= \mathbf{P}(\min\{1-X, X\} \le v) = \mathbf{P}(1-X \le v \text{ or } X \le v) \\ &= \mathbf{P}(1-X \le v) + \mathbf{P}(X \le v) = \mathbf{P}(1-v \le X) + \mathbf{P}(X \le v) \\ &= 1-(1-v)+v = 2v.\end{aligned}$$

2nd answer. For $v \in (0, 1/2)$,

$$\begin{aligned}\mathbf{P}(V \le v) &= \mathbf{P}(\min\{1-X, X\} \le v) \\ &= \mathbf{P}(1-X \le X, 1-X \le v) + \mathbf{P}(X \le 1-X, X \le v) \\ &= \mathbf{P}(1-v \le X) + \mathbf{P}(X \le v) = 2v.\end{aligned}$$

3rd answer. For $v \in (0, 1/2)$,

$$\begin{aligned}\mathbf{P}(V \le v) &= 1-\mathbf{P}(V > v) = 1-\mathbf{P}(X > v, 1-X > v) \\ &= 1-\mathbf{P}(v < X < 1-v) = 1-(1-2v) = 2v.\end{aligned}$$

Hence $F_V(v) = 2v$ if $0 < v \le 1/2$, $F_V(v) = 1$ if $v \ge 1$, and zero otherwise. Then $f_V(v) = F'(v) = 2$ if $v \in (0, 1/2)$, and zero otherwise. V is uniform in $(0, 1/2)$.

b) Let X, Y be independent uniform in $(0, 1)$, and $V = \frac{Y}{X}$. What is the set of possible values of V? Find the df and pdf of $V = \frac{Y}{X}$. Hint. Draw a picture and determine the areas under $\frac{y}{x} \le v$.

Answer. The range of $V = Y/X$ is $(0, \infty)$. Drawing a picture and finding some triangle areas, we see that for $v \in (0, 1)$, the df

$$F(v) = \mathbf{P}(V \le v) = \mathbf{P}(Y \le vX) = \frac{1}{2}v,$$

For $v \ge 1$,

$$F(v) = 1 - \frac{1}{2v},$$

$F(v) = 0$ if $v \le 0$. The pdf $f(v) = F'(v) = \frac{1}{2}$if $v \in (0, 1)$, $f(v) = F'(v) = \frac{1}{2v^2}$ if $v \ge 1$, (zero otherwise).

5. (St. Petersburg paradox) You pay $\$2^8 = \$256$ to enter and play the following game: A fair coin is tossed until H shows up, and if X is the number of tosses that was needed, you are paid $\$2^X$. For instance, if $X = 10$, then you are paid $\$2^{10}$).

Your win/loss $W = 2^X - 256$. What is $\mathbf{E}(W)$? What is the probability that $W \ge 0$?

Answer. We know that X is geometric r.v.: $\mathbf{P}(X = k) = 2^{-k}, k = 1, 2, \ldots$. Hence

$$\begin{aligned}\mathbf{E}(W) &= \mathbf{E}\left(2^X\right) - 256 \\ &= \sum_{k=1}^{\infty} 2^k 2^{-k} - 256 = \sum_{k=1}^{\infty} 1 - 256 = +\infty.\end{aligned}$$

On the other hand, "we do not lose" = $\{W \ge 0\} = \left\{2^X \ge 2^8\right\} = \{X \ge 8\}$, and

$$\mathbf{P}(W \ge 0) = \mathbf{P}(X \ge 8) = \sum_{k=8}^{\infty} 2^{-k} = 2^{-7} = \frac{1}{128}.$$

12 hw12

1. The joint mgf of a random vector $X = (X_1, \ldots, X_d)$ is defined as

$$M_X(t) = \mathbf{E}\left[e^{t_1 X_1 + \ldots t_d X_d}\right] = \mathbf{E}\left[e^{tX'}\right], t = (t_1, \ldots, t_d) \in \mathbf{R}^d,$$

where tX' is the product of the row vector $t = (t_1, \ldots, t_d)$ and the column vector X', the transpose of random row vector $X = (X_1, \ldots, X_d)$.

(i) Let $Z = (Z_1, \ldots, Z_d)$ with independent $Z_i \sim N(0, 1)$. Show that

$$M_Z(t) = \exp\left\{\frac{1}{2}tt'\right\} = \exp\left\{\frac{1}{2}\left(t_1^2 + \ldots + t_d^2\right)\right\}, t = (t_1, \ldots, t_d) \in \mathbf{R}^d.$$

Answer. By definition, and using independence,

$$\begin{aligned} M_Z(t) &= \mathbf{E}\left[e^{t_1 Z_1 + \ldots t_d Z_d}\right] = \mathbf{E}\left[e^{tZ'}\right] \\ &= \mathbf{E}\left(e^{t_1 Z_1} \ldots e^{t_d Z_d}\right) = M_{Z_1}(t_1) \ldots M_{Z_d}(t_d) \\ &= e^{t_1^2/2} \ldots e^{t_d^2/2} = \exp\left\{\frac{1}{2}\left(t_1^2 + \ldots + t_d^2\right)\right\} = \exp\left\{\frac{1}{2}tt'\right\}. \end{aligned}$$

(ii) Let $X = (X_1, \ldots, X_d)$ be multivariate normal $N(\mu, B)$. Show that

$$\begin{aligned} M_X(t) &= \exp\left\{t\mu' + \frac{1}{2}tBt'\right\} \\ &= \exp\left\{\sum_{j=1}^{d} t_j \mu_j + \frac{1}{2}\sum_{k,j=1}^{d} b_{kj} t_k t_j\right\}, t = (t_1, \ldots, t_d) \in \mathbf{R}^d. \end{aligned}$$

Hint. We have $B = D'D$ for some $d \times d$ -matrix D, and $X = ZD + \mu$, where $Z = (Z_1, \ldots, Z_d)$ is r. vector of independent $Z_i \sim N(0, 1)$: use definition of M_X and part a).

Answer. Since $B = D'D$ for some $d \times d$ -matrix D, and $X = ZD + \mu$, where $Z = (Z_1, \ldots, Z_d)$ is r. vector of independent $Z_i \sim N(0, 1)$, we have for $t = (t_1, \ldots, t_d)$, by part a),

$$\begin{aligned} tX' &= t\mu' + t(ZD)' = t\mu' + tD'Z', \\ M_X(t) &= \mathbf{E}\left(e^{tX'}\right) = e^{t\mu'}\mathbf{E}\left(e^{tD'Z'}\right) = e^{t\mu'} M_Z\left(tD'\right) \\ &= \exp\left\{t\mu' + \frac{1}{2}tD'\left(tD'\right)'\right\} = \exp\left\{it\mu' + \frac{1}{2}tD'Dt'\right\} \\ &= \exp\left\{t\mu' + \frac{1}{2}tBt'\right\}. \end{aligned}$$

2. Consider a branching process with immigration: each generation is supplemented by an "immigrant" with probability p. This means that the size Z_n of the n-th generation satisfies

$$Z_n = I_n + \sum_{k=1}^{Z_{n-1}} X_k,$$

where $I_n = 1$ with probability p and $I_n = 0$ otherwise; the number of children X_k of the kth person in the generation $n-1$ are independent identically distributed with generating function $G(s)$ and mean μ. We assume that Z_{n-1}, I_n and X_k are independent. Let $G_n(s) = G_{Z_n}(s)$ and $\mu_n = \mathbf{E}(Z_n)$.

(a) Show that $G_n(s) = [ps + (1-p)]\, G_{n-1}(G(s))$.

Answer. Since I_n and $\sum_{k=1}^{Z_{n-1}} X_k$ are independent, we have

$$\begin{aligned} G_n(s) &= \mathbf{E}\left(s^{Z_n}\right) = \mathbf{E}\left(s^{I_n + \sum_{k=1}^{Z_{n-1}} X_k}\right) = \mathbf{E}\left(s^{I_n} s^{\sum_{k=1}^{Z_{n-1}} X_k}\right) \\ &= \mathbf{E}\left(s^{I_n}\right)\mathbf{E}\left(s^{\sum_{k=1}^{Z_{n-1}} X_k}\right) = (ps+q)\,\mathbf{E}\left(s^{\sum_{k=1}^{Z_{n-1}} X_k}\right) = (ps+q)\, G_{n-1}(G(s)), \end{aligned}$$

by a theorem we proved in class (Thm1, note of 11/16) applied to $\mathbf{E}\left(s^{\sum_{k=1}^{Z_{n-1}} X_k}\right) = G_{\sum_{k=1}^{Z_{n-1}} X_k}(s)$.

Alternatively, conditioning on Z_{n-1} and using independence of X_k,

$$\mathbf{E}\left(s^{\sum_{k=1}^{Z_{n-1}} X_k}\right) = \mathbf{E}\left\{\mathbf{E}\left(s^{\sum_{k=1}^{Z_{n-1}} X_k} | Z_{n-1}\right)\right\},$$

and

$$\begin{aligned} \mathbf{E}\left(s^{\sum_{k=1}^{Z_{n-1}} X_k} | Z_{n-1} = j\right) &= \mathbf{E}\left(s^{\sum_{k=1}^{j} X_k} | Z_{n-1} = j\right) = \mathbf{E}\left(s^{\sum_{k=1}^{j} X_k}\right) = G_X(s)^j, \\ \mathbf{E}\left\{\mathbf{E}\left(s^{\sum_{k=1}^{Z_{n-1}} X_k} | Z_{n-1}\right)\right\} &= G_X(s)^{Z_{n-1}}, \\ \mathbf{E}\left(s^{\sum_{k=1}^{Z_{n-1}} X_k}\right) &= \mathbf{E}\left[G_X(s)^{Z_{n-1}}\right] = G_{n-1}(G_X(s)) = G_{n-1}(G(s)). \end{aligned}$$

(b) Show that $\mu_n = p + \mu\mu_{n-1}$;

Answer. Recall for any r.v. V with values in $\{0, 1, \ldots\}$, $G_V(1) = 1, G_V'(1) = \mathbf{E}(V)$. Since $\mu_n = G_n'(1)$, we differentiate

$$\frac{d}{ds} G_n(s) = pG_{n-1}(G(s)) + (q + sp)\, G_{n-1}'(G(s))\, G'(s)$$

and, by chain rule,

$$\begin{aligned} \mu_n &= \frac{d}{ds} G_n(1) = pG_{n-1}(G(1)) + G_{n-1}'(G(1))\, G'(1) \\ &= pG_{n-1}(1) + G_{n-1}'(1)\, G'(1) = p + \mu_n \mu. \end{aligned}$$

Or other way,

$$\mathbf{E}(Z_n) = \mathbf{E}(I_n) + \mathbf{E}\left(\sum_{k=1}^{Z_{n-1}} X_k\right) = p + \mathbf{E}\left[\mathbf{E}\left(\sum_{k=1}^{Z_{n-1}} X_k | Z_{n-1}\right)\right],$$

and

$$\begin{aligned} \mathbf{E}\left[\mathbf{E}\left(\sum_{k=1}^{Z_{n-1}} X_k | Z_{n-1} = j\right)\right] &= \mathbf{E}\left[\mathbf{E}\left(\sum_{k=1}^{j} X_k | Z_{n-1} = j\right)\right] = \mathbf{E}\left[\mathbf{E}\left(\sum_{k=1}^{j} X_k\right)\right] = j\mu, \\ \mathbf{E}\left[\mathbf{E}\left(\sum_{k=1}^{Z_{n-1}} X_k | Z_{n-1}\right)\right] &= \mu Z_{n-1}, \mathbf{E}(Z_n) = p + \mu\mathbf{E}(Z_{n-1}) = p + \mu\mu_{n-1}. \end{aligned}$$

(c) Find μ_n and $\lim_{n\to\infty}\mu_n$.

Answer. "Going down" in (b), we find that (with $\mu_0 = \mathbf{E}(Z_0)$),

$$\begin{aligned}\mu_n &= p + \mu\mu_{n-1} = p + \mu(p + \mu\mu_{n-2}) = p(1+\mu) + \mu^2\mu_{n-2} = \dots \\ &= p\left(1 + \mu + \dots + \mu^{n-1}\right) + \mu^n\mu_0 = p\sum_{j=0}^{n-1}\mu^j + \mu^n\mu_0,\end{aligned}$$

and

$$\mu_n \to p\sum_{j=0}^{\infty}\mu^j = \frac{p}{1-\mu} \text{ if } \mu < 1,$$

and $\mu_n = np + \mu_0 \to \infty$ if $\mu = 1$, $\mu_n \to \infty$ if $\mu > 1$.

3. a) A sequence of biased coins is flipped; the chance that the kth coin shows a head is U_k, where U_k is a random variable taking values in $(0,1)$. Let X_n be the number of heads after n flips. Does X_n obey the central limit theorem when the U_k are independent and identically distributed?

Answer. Let

$$V_i = \begin{cases} 1 & \text{if } H \text{ in the } i\text{th flip} \\ 0 & \text{otherwise.}\end{cases}$$

Then $V_i \sim$Bernoulli(U_i). Since U_i are independent, V_i are independent Bernoulli with parameter

$$\mathbf{P}(V_i = 1) = \mathbf{E}\{\mathbf{P}(V_i = 1|U_i)\} = \mathbf{E}(U_i) =: p_0 \in (0,1),$$

that is $X_n = \sum_{i=1}^n V_i$ is Binomial(n, p_0). So, CLT holds for it:

$$\frac{X_n - np_0}{\sqrt{np_0(1-p_0)}} \xrightarrow{D} Z \sim N(0,1).$$

Comment. If U_k are independent uniform in $(0,1)$, then $p_0 = 1/2$: it is like tossing a sequence of fair coins.

b) Let X_n be binomial(n, U), where U is uniform in $(0,1)$ (look at #1 of hw10). Show that

$$\frac{X_n}{n+1} \xrightarrow{D} U,$$

that is CLT does not hold for X_n.

Answer. We find

$$\begin{aligned}\Phi_U(t) &= \mathbf{E}\left(e^{itU}\right) = \int_0^1 e^{itu}du \\ &= \frac{e^{itu}}{it}|_0^1 = \frac{e^{it}-1}{it} \text{ if } t \neq 0, \Phi_U(0) = 1.\end{aligned}$$

Now,

$$\begin{aligned}\Phi_{X_n}(t) &= \mathbf{E}\left(e^{itX_n}\right) = \mathbf{E}\left[\mathbf{E}\left(e^{itX_n}|U\right)\right] = \mathbf{E}\left[\left(1 - U + e^{it}U\right)^n\right] \\ &= \int_0^1 \left(1 + (e^{it}-1)u\right)^n du = \frac{\left(1 + (e^{it}-1)u\right)^{n+1}}{(e^{it}-1)(n+1)}|_0^1 \\ &= \frac{e^{it(n+1)}-1}{(n+1)(e^{it}-1)}, t \neq 0, \Phi_{X_n}(0) = 1.\end{aligned}$$

Hence the characteristic function

$$\begin{aligned}
\Phi_{\frac{X_n}{n+1}}(t) &= \Phi_{X_n}\left(\frac{t}{n+1}\right) = \frac{e^{i\frac{t}{n+1}(n+1)}-1}{(n+1)\left(e^{i\frac{t}{n+1}}-1\right)} = \frac{e^{it}-1}{(n+1)\left(e^{i\frac{t}{n+1}}-1\right)} \to \frac{e^{it}-1}{it} \\
&= \Phi_U(t),
\end{aligned}$$

as $n \to \infty$, because by definition of the derivative or L'Hospital,

$$\lim_n (n+1)\left(e^{i\frac{t}{n+1}}-1\right) = \lim_n it\frac{e^{i\frac{t}{n+1}}-1}{\frac{it}{n+1}} = it\lim_{z\to 0}\frac{e^z-1}{z} = it.$$

Since the limit is uniform, CLT does not hold.

Comment. Assume a machine makes coins whose heads probability is U, uniform r.v. in $(0,1)$. If we find a coin made by that machine and flip it n times, then X_n, the number of heads in n flips, is binomial(n,U). Hence

$$X_n = Y_1 + \ldots + Y_n,$$

where $Y_i \sim$Bernoulli(p) with

$$p = \mathbf{P}(Y_i = 1) = \mathbf{E}[\mathbf{P}(Y_i = 1|U)] = \mathbf{E}(U) = \frac{1}{2}.$$

Although Y_i are identically distributed, Y_i are not independent.

4. a) Let H_n be number of H in n independent tosses of a p-coin. Apply CLT to approximate $\mathbf{P}\left(a < \frac{H_n}{n} < b\right)$, $0 < a < b < 1$, for large n.

Answer. Let X_i be the number of heads in the ith toss. Then $X_1, X_2, \ldots$ are independent Bernoulli(p), $\mathbf{E}(X_i) = p$, var$(X_i) = pq$, and

$$H_n = \sum_{i=1}^n X_i.$$

By CLT, for large n,

$$\begin{aligned}
\mathbf{P}\left(a < \frac{H_n}{n} = \bar{X}_n < b\right) &\approx \mathbf{P}\left(\frac{a-p}{\sqrt{\frac{p(1-p)}{n}}} < Z < \frac{b-p}{\sqrt{\frac{p(1-p)}{n}}}\right) \\
&= F_Z\left(\frac{b-p}{\sqrt{\frac{p(1-p)}{n}}}\right) - F_Z\left(\frac{a-p}{\sqrt{\frac{p(1-p)}{n}}}\right),
\end{aligned}$$

where F_Z is the df of a standard normal.

Comment. Since H_n is binomial(n,p), for $p = q = 1/2$, we have for large n,

$$\sum_{a<\frac{j}{n}<b}\binom{n}{j} \approx 2^n\frac{1}{\sqrt{2\pi}}\int_{\sqrt{n}(2a-1)}^{\sqrt{n}(2b-1)} e^{-x^2/2}dx.$$

b) Let Y be Poisson(λ). Apply continuity theorem to show that

$$\frac{Y-\lambda}{\sqrt{\lambda}} \xrightarrow{D} Z \sim N(0,1)$$

as $\lambda \to \infty$. Hint. Recall $\mathbf{E}(Y) = \text{Var}(Y) = \lambda$. Find characteristic function of $V = \frac{Y-\lambda}{\sqrt{\lambda}}$.

Comment. For large λ, Y is approximately $N(\lambda, \lambda)$.

Answer. By continuity theorem we need to show that for all t,

$$\phi_{\frac{Y-\lambda}{\sqrt{\lambda}}}(t) \to \phi_Z(t) = e^{-t^2/2}.$$

Since $\phi_Y(t) = \exp\left\{\lambda\left(e^{it}-1\right)\right\}$, we find

$$\begin{aligned}
\phi_{\frac{Y-\lambda}{\sqrt{\lambda}}}(t) &= \phi_{\frac{1}{\sqrt{\lambda}}Y-\sqrt{\lambda}}(t) = \exp\left\{-i\sqrt{\lambda}t\right\}\phi_{\frac{1}{\sqrt{\lambda}}Y}(t) = \exp\left\{-i\sqrt{\lambda}t\right\}\phi_Y\left(\frac{t}{\sqrt{\lambda}}\right) \\
&= \exp\left\{\lambda\left(e^{i\frac{t}{\sqrt{\lambda}}}-1-i\frac{t}{\sqrt{\lambda}}\right)\right\}.
\end{aligned}$$

Now,

$$\lambda\left(e^{i\frac{t}{\sqrt{\lambda}}}-1-i\frac{t}{\sqrt{\lambda}}\right) = \left(-\frac{t^2}{2}\right)\frac{e^{i\frac{t}{\sqrt{\lambda}}}-1-i\frac{t}{\sqrt{\lambda}}}{\frac{1}{2}\left(\frac{it}{\sqrt{\lambda}}\right)^2}$$

and, denoting $z = it/\sqrt{\lambda} \to 0$ as $\lambda \to \infty$, by L'Hospital,

$$\frac{e^{i\frac{t}{\sqrt{\lambda}}}-1-i\frac{t}{\sqrt{\lambda}}}{\frac{1}{2}\left(\frac{it}{\sqrt{\lambda}}\right)^2} = \frac{e^z-1-z}{\frac{1}{2}z^2} \to 1 \text{ as } z \to 0.$$

Hence

$$\phi_{\frac{Y-\lambda}{\sqrt{\lambda}}}(t) \to e^{-t^2/2} = \phi_Z(t)$$

as $\lambda \to \infty$. The claim follows by continuity theorem..

5. Let $X_k, k \geq 1$, be i.i.d. random variables with mean 1 and variance 1. Show that the limit

$$\lim_{n\to\infty}\frac{\sum_{k=1}^n X_k}{\sum_{k=1}^n X_k^2}$$

exists in an appropriate sense, and identify the limit.

Answer. By strong law of large numbers,

$$\begin{aligned}
\frac{\sum_{k=1}^n X_k}{n} &\to \mathbf{E}(X) = 1, \\
\frac{\sum_{k=1}^n X_k^2}{n} &\to \mathbf{E}\left(X^2\right) = \text{Var}(X) + (\mathbf{E}(X))^2 = 1+1 = 2
\end{aligned}$$

with probability 1 as $n \to \infty$. Hence

$$\frac{\sum_{k=1}^n X_k}{\sum_{k=1}^n X_k^2} = \frac{\left(\sum_{k=1}^n X_k\right)/n}{\left(\sum_{k=1}^n X_k^2\right)/n} \to \frac{1}{2}$$

with probability 1 as $n \to \infty$.