Intro to Probability
Solution Notes for Example Sheet 3

1 Sum of distributions

Extended Note 1 [Computing the random variables] Consider two discrete independent random variables X and Y with pmfs f_X and f_Y. Then, we want to compute the pmf for the random variable $Z = X + Y$.
In order to do this we sum up the probabilities for all ways of making the sum z.

$$f_Z(z) = \mathbf{Pr}[Z = z] = \sum_{k=-\infty}^{\infty} \mathbf{Pr}[X = k, Y = z - k] = \sum_{k=-\infty}^{\infty} \mathbf{Pr}[X = k] \cdot \mathbf{Pr}[Y = z - k]$$
$$= \sum_{k=-\infty}^{\infty} f_X(k) \cdot f_Y(z - k).$$

Using this formula we can compute the pmf for Z.
For continuous random variables X and Y with pdfs f_X and f_Y, the formula becomes

$$f_Z(z) = \int_{k=-\infty}^{\infty} f_X(k) \cdot f_Y(z - k)\, dk.$$

This type of summation is also known as *convolution* and it is used in several places, like signal processing, computer vision or efficient computation (see this video if you would like to learn more).

Exercise 1 [Sum of Poisson r.vs.] Consider two independent Poisson r.vs. $X \sim \mathsf{Poi}(\mu)$ and $Y \sim \mathsf{Poi}(\lambda)$. Show that $Z = X + Y \sim \mathsf{Poi}(\mu + \lambda)$.

(Answer) Using the above formula we have that

$$\begin{aligned}
f_Z(z) &= \sum_{k=-\infty}^{\infty} f_X(k) \cdot f_Y(z - k) \\
&= \sum_{k=0}^{z} \frac{e^{-\mu} \cdot \mu^k}{k!} \cdot \frac{e^{-\lambda} \cdot \lambda^{z-k}}{(z-k)!} \\
&= e^{-\mu-\lambda} \sum_{k=0}^{z} \frac{\mu^k}{k!} \cdot \frac{\lambda^{z-k}}{(z-k)!} \\
&= \frac{1}{z!} \cdot e^{-\mu-\lambda} \sum_{k=0}^{z} \frac{z!}{k!(z-k)!} \cdot \mu^k \cdot \lambda^{z-k} \\
&= \frac{1}{z!} \cdot e^{-\mu-\lambda} \sum_{k=0}^{z} \binom{z}{k} \cdot \mu^k \cdot \lambda^{z-k} \\
&= \frac{e^{-\mu-\lambda} \cdot (\mu + \lambda)^z}{z!},
\end{aligned}$$

using in the last step the Binomial sum formula, i.e., $(x + y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$. Therefore, $Z \sim \mathsf{Poi}(\mu + \lambda)$.

Exercise 2 [Sum of uniform distributions] Consider three independent uniform distributions $X_1, X_2, X_3 \in \mathcal{U}[0, 1]$.

(a) Determine the pdf for $S_2 = X_1 + X_2$.

(b) Determine the pdf for $S_3 = X_1 + X_2 + X_3$.

(Answer)

(a) Using the formula, we have that

$$f_{S_2}(z) = \int_{k=-\infty}^{\infty} f_{X_1}(k) \cdot f_{X_2}(z-k)\, dk = \int_{k=0}^{1} f_{X_2}(z-k)\, dk.$$

Case A [$z \in [0,1]$]: Here k can be in $[0, z]$, so

$$f_{S_2}(z) = \int_{k=0}^{z} 1\, dk = k\Big|_0^z = z.$$

Case B [$z \in [1,2]$]: Here k can be in $[z-1, 1]$, so

$$f_{S_2}(z) = \int_{k=z-1}^{1} 1\, dk = k\Big|_{z-1}^{1} = 2 - z.$$

Combining these cases, we deduce that

$$f_{S_2}(z) = \begin{cases} z & \text{if } z \in [0,1] \\ 2-z & \text{if } z \in (1,2] \\ 0 & \text{otherwise.} \end{cases}$$

(b) For three random variables, we are going to use the pdf for $S_2 = X_1 + X_2$,

$$f_{S_3}(z) = \int_{k=-\infty}^{\infty} f_{X_1+X_2}(z-k) \cdot f_{X_3}(k)\, dk = \int_{k=0}^{1} f_{X_1+X_2}(z-k)\, dk.$$

We now consider three cases based on the value of z:
Case A [$z \in [0,1]$]: Here k can be in $[0, z]$, so

$$f_{S_3}(z) = \int_{k=0}^{z} (z-k)\, dk = (zk - k^2/2)\Big|_0^z = \frac{z^2}{2}.$$

Case B [$z \in [1,2]$]: Here k can be in $[0,1]$ and we break the integral depending on whether $z-1=1$ or not. So,

$$\begin{aligned} f_{S_3}(z) &= \int_{k=0}^{z-1} f_{X_1+X_2}(z-k)\, dk + \int_{k=z-1}^{1} f_{X_1+X_2}(z-k)\, dk \\ &= \int_{k=0}^{z-1} (2-z+k)\, dk + \int_{k=z-1}^{1} (z-k)\, dk \\ &= (2-z)\cdot(z-1) + (z-1)^2/2 + z - 1/2 - (z-1)\cdot z + (z-1)^2/2 \\ &= -z^2 + 3z - \frac{3}{2}. \end{aligned}$$

Case C [$z \in [2,3]$]: This case is symmetric to Case A. Therefore,

$$f_{S_3}(z) = \frac{1}{2} \cdot (3-z)^2.$$

Exercise 3 Given the following pmf for random variables X and Y, compute the pmf for $Z = X + Y$.

	1	2	3	4
$\mathbf{Pr}[X = x]$	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{1}{6}$	$\frac{2}{6}$
$\mathbf{Pr}[Y = y]$	0	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$

(Answer)

	1	2	3	4	5	6	7	8
$\mathbf{Pr}[X+Y = z]$	0	0	$\frac{1}{24}$	$\frac{2+2}{24}$	$\frac{1+4+1}{24}$	$\frac{2+2+2}{24}$	$\frac{4+1}{24}$	$\frac{2}{24}$

Question 1: How would you compute the pmf for $X - Y$?
Question 2: How would you recover the marginal distributions given the pmf for $X + Y$ and $X - Y$?

2 Minimum/Maximum of random variables

Extended Note 2 [Computing the distribution function] Given two independent random variables X and Y with cummulative distribution functions F_X and F_Y, we want to compute the cummulative distribution function F_Z for $Z = \max\{X, Y\}$.
The main observation is to see that $\max\{X, Y\} \leq z$ iff both $X \leq z$ and $y \leq z$ (*Why?*). Then, we obtain

$$F_Z(z) = \mathbf{Pr}[Z \leq z] = \mathbf{Pr}[\max\{X, Y\} \leq z] = \mathbf{Pr}[X \leq z, Y \leq z] = \mathbf{Pr}[X \leq z] \cdot \mathbf{Pr}[Y \leq z]$$
$$= F_X(z) \cdot F_Y(z).$$

Similarly for $Z = \min\{X, Y\}$ we have that

$$F_Z(z) = \mathbf{Pr}[Z \leq z] = 1 - \mathbf{Pr}[Z > z] = 1 - \mathbf{Pr}[\min\{X, Y\} > z] = 1 - \mathbf{Pr}[X > z, Y > z]$$
$$= 1 - \mathbf{Pr}[X > z] \cdot \mathbf{Pr}[Y > z] = 1 - (1 - F_X(z)) \cdot (1 - F_Y(z)).$$

Exercise 4 [Minimum of uniform r.vs.] Consider n independent uniform random variables $X_1, \ldots, X_n \sim \mathcal{U}[0, 1]$.

(a) Determine the cummulative distribution function for $Z = \max\{X_1, \ldots, X_n\}$.

(b) Determine the probability density function for Z.

(c) Determine the expectation for Z.

(Answer)

(a) Using the above technique for n random variables we have that for any $z \in [0, 1]$,

$$\begin{aligned} F_Z(z) &= \mathbf{Pr}[Z \leq z] = \mathbf{Pr}[\max\{X_1, \ldots, X_n\} \leq z] \\ &= \mathbf{Pr}[X_1 \leq z] \cdot \ldots \cdot \mathbf{Pr}[X_n \leq z] \\ &= z \cdot \ldots \cdot z \\ &= z^n. \end{aligned}$$

Therefore,

$$F_Z(z) = \begin{cases} 0 & \text{if } z < 0 \\ z^n & \text{if } z \in [0, 1] \\ 1 & \text{otherwise.} \end{cases}$$

(b) By differentiating, we get the pdf for Z:

$$f_Z(z) = \frac{d}{dz} F_Z(z) = \begin{cases} 0 & \text{if } z < 0 \\ n \cdot z^{n-1} & \text{if } z \in [0, 1] \\ 0 & \text{otherwise.} \end{cases}$$

(c) For the expectation of Z, we have that

$$\begin{aligned} \mathbf{E}[Z] &= \int_{z=0}^{1} f_Z(z) \cdot z \, dz \\ &= \int_{z=0}^{1} n \cdot z^{n-1} \cdot z \, dz \\ &= \int_{z=0}^{1} n \cdot z^n \, dz \\ &= n \cdot \left. \frac{z^{n+1}}{n+1} \right|_0^1 \\ &= \frac{n}{n+1}. \end{aligned}$$

Exercise 5 [Minimum of Exponential r.vs.] Consider two independent Exponential r.vs. $X \sim \mathsf{Exp}(\lambda)$ and $Y \sim \mathsf{Exp}(\mu)$. Find the cummulative distribution of $Z = \min\{X, Y\}$.

(Answer) Recall that the cdf of an Exponential r.v. is given by

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{otherwise,} \end{cases} \quad \text{and} \quad F_Y(y) = \begin{cases} 1 - e^{-\mu y} & \text{if } y \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, for any $z \geq 0$, we have that

$$F_Z(z) = 1 - (1 - F_X(z)) \cdot (1 - F_Y(z)) = 1 - e^{-\lambda z} \cdot e^{-\mu z} = 1 - e^{-(\lambda+\mu)\cdot z}.$$

Therefore, Z follows $\mathsf{Exp}(\lambda + \mu)$.

Exercise 6 [Minimum of geometric r.vs.] Consider two independent Geometric r.vs. $X \sim \mathsf{Geom}(p)$ and $Y \sim \mathsf{Geom}(q)$. Find the cummulative distribution of $Z = \min\{X, Y\}$.

(Answer) Recall that the cdf of a Geometric r.v. is given by

$$F_X(x) = \begin{cases} 1 - (1-p)^x & \text{if } x \geq 0 \\ 0 & \text{otherwise,} \end{cases} \quad \text{and} \quad F_Y(y) = \begin{cases} 1 - (1-q)^y & \text{if } y \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Using the formula above, we have that

$$F_Z(z) = 1 - (1 - F_X(z)) \cdot (1 - F_Y(z)) = 1 - (1-p)^z \cdot (1-q)^z = 1 - (1 - p - q + pq)^z.$$

Therefore, Z follows $\mathsf{Geom}(p + q - pq)$.

3 Joint and marginal distributions

Exercise 7 The joint probability density function of X and Y is given by

$$f(x, y) = c \cdot (y^2 - x^2) \cdot e^{-y}, \quad -y \leq x \leq y, 0 < y < \infty.$$

(a) Find c.

(b) Find the marginal densities of X and Y.

(c) Find $\mathbf{E}[X]$.

(Answer) See solution to problem 9 here.

Exercise 8 The joint probability density function of X and Y is given by

$$f(x, y) = e^{-x-y}, \quad 0 \leq x < \infty, 0 \leq y < \infty.$$

(a) Find $\mathbf{Pr}[X < Y]$.

(b) Find $\mathbf{Pr}[X < a]$.

(Answer) See solution to problem 10 here.

Exercise 9 The joint probability density function of X and Y is given by

$$f(x, y) = 12xy(1 - x), \quad 0 < x < 1, 0 < y < 1.$$

(a) Are X and Y independent?

(b) Find $\mathbf{E}[X]$ and $\mathbf{E}[Y]$.

(c) Find $\mathbf{Var}[X]$ and $\mathbf{Var}[Y]$.

(Answer) See solution to problem 23 here.

Exercise 10 Suppose that X and Y have a discrete joint distribution for which the joint PMF is defined as follows:

$$f(x,y) = \begin{cases} c|x+y|, & x=-1,0,1 \text{ and } y=-1,0,1 \\ 0, & \textit{otherwise}. \end{cases}$$

Determine:

(a) Determine c.

(b) Determine $\mathbf{Pr}\left[\, X=0, Y=1 \,\right]$ and $\mathbf{Pr}\left[\, X=1 \,\right]$.

(c) Determine $\mathbf{Pr}\left[\, |X-Y|<1 \,\right]$.

(Answer) See solution to problem 1 here.

Further Reading 1 [Further exercises] You can find more exercises with solutions here and here.

4 Computing the variance

Exercise 11 [Hats] There are n people taking their hats randomly. Let N be the total number of people that got the correct hat back.

(a) Show that $\mathbf{E}\left[\, N \,\right] = 1$.

(b) Show that $\mathbf{Var}\left[\, N \,\right] = 1$.

(c) Use Chebyshev's inequality to deduce bounds on N.

(Answer)

(a) Let X_i be the indicator of the event $\mathcal{E}_i = \{\text{Person } i \text{ got their hat back}\}$. Then

$$\mathbf{E}\left[\, X_i \,\right] = \mathbf{Pr}\left[\, \mathcal{E}_i \,\right] = \frac{1}{n}.$$

By linearity of expectation, we have that

$$\mathbf{E}\left[\, N \,\right] = \mathbf{E}\left[\sum_{i=1}^n X_i \right] = \sum_{i=1}^n \mathbf{E}\left[\, X_i \,\right] = n \cdot \frac{1}{n} = 1.$$

(b) Using the formula for the variance

$$\begin{aligned} \mathbf{Var}\left[\, X \,\right] &= \mathbf{E}\left[\, X^2 \,\right] - (\mathbf{E}\left[\, X \,\right])^2 \\ &= \mathbf{E}\left[\left(\sum_{i=1}^n X_i \right)^2 \right] - 1 \\ &= \sum_{i=1}^n \sum_{j=1}^n \mathbf{E}\left[\, X_i X_j \,\right]. \end{aligned}$$

We distinguish two cases for $\mathbf{E}\left[\, X_i X_j \,\right]$:

- **Case [$i=j$]:** Since $X_i \in \{0,1\}$ we have that $X_i = X_i^2$, so

$$\mathbf{E}\left[\, X_i^2 \,\right] = \mathbf{E}\left[\, X_i \,\right] = \frac{1}{n}.$$

- **Case [$i \neq j$]:** This one is a bit more involved:

$$\begin{aligned} \mathbf{E}\left[\, X_i X_j \,\right] &= \mathbf{Pr}\left[\, X_i X_j = 1 \,\right] = \mathbf{Pr}\left[\, X_i X_j = 1 \mid X_i = 1 \,\right] \cdot \mathbf{Pr}\left[\, X_i = 1 \,\right] \\ &= \mathbf{Pr}\left[\, X_j = 1 \mid X_i = 1 \,\right] \cdot \mathbf{Pr}\left[\, X_i = 1 \,\right] \\ &= \frac{1}{n-1} \cdot \frac{1}{n}, \end{aligned}$$

since given that person i got the correct hat back, there are $n-1$ remaining items in $n-1$ slots, so the probability that j also got the correct hat back is $1/(n-1)$.

By combining the two cases, we have that

$$\mathbf{Var}\left[X \right] = n \cdot \frac{1}{n} + n \cdot (n-1) \cdot \frac{1}{n \cdot (n-1)} - 1 = 1.$$

Exercise 12 [Max-Cut] In Part IA Algorithms, you saw the Min-Cut problem, where given a graph $G=(V,E)$ the goal is to find a subset $S \subseteq V$ such that the number of the edges crossing S and $V \setminus S$ is minimised. In this exercise, we will look at the problem of *maximising* the number edges crossing the cut.

Consider the algorithm that goes through the vertices one by one and adds it to S independently with probability $1/2$.

(a) Show that the exected size C of the cut produced is $|E|/2$. Argue that this is within a factor 2 of the optimal.

(b) Compute the $\mathbf{Var}\left[C \right]$.

(c) Use Chebyshev's inequality to deduce bounds on C.

(Answer)

(a) For each edge $e \in E$, let X_e be the indicator of the event $\mathcal{E}_e = \{\text{edge } e \text{ crosses the cut}\}$. Then, for edge $e=(u,v)$, then

$$\mathbf{E}\left[X_e \right] = \mathbf{Pr}\left[\mathcal{E}_e \right] = \mathbf{Pr}\left[u \in S, v \notin S \right] + \mathbf{Pr}\left[u \notin S, v \in S \right] = \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{2}.$$

Therefore, by linearity of expectation

$$\mathbf{E}\left[C \right] = \mathbf{E}\left[\sum_{e \in E} X_e \right] = \sum_{e \in E} \mathbf{E}\left[X_e \right] = |E|/2.$$

The maximum cut can have at most $|E|$ edges, therefore, this simple algorithm gives a 2-approximation for max-cut in expectation.

(b) Using the formula for the variance

$$\begin{aligned}
\mathbf{Var}\left[X \right] &= \mathbf{E}\left[X^2 \right] - (\mathbf{E}\left[X \right])^2 \\
&= \mathbf{E}\left[\left(\sum_{i=1}^{n} X_i \right)^2 \right] - |E|^2/4 \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} \mathbf{E}\left[X_i X_j \right] - |E|^2/4.
\end{aligned}$$

We distinguish two cases for $\mathbf{E}\left[X_i X_j \right]$:

- **Case $[i=j]$:** Since $X_i \in \{0,1\}$ we have that $X_i = X_i^2$, so

$$\mathbf{E}\left[X_i^2 \right] = \mathbf{E}\left[X_i \right] = \frac{1}{2}.$$

- **Case $[i \neq j]$:** This one is a bit more involved

$$\begin{aligned}
\mathbf{E}\left[X_i X_j \right] = \mathbf{Pr}\left[X_i X_j = 1 \right] &= \mathbf{Pr}\left[X_i X_j = 1 \mid X_i = 1 \right] \cdot \mathbf{Pr}\left[X_i = 1 \right] \\
&= \mathbf{Pr}\left[X_j = 1 \mid X_i = 1 \right] \cdot \mathbf{Pr}\left[X_i = 1 \right].
\end{aligned}$$

Now we consider further subcases based on the number of common vertices between edges $i = (u-1, v_1)$ and $j = (u_2, v_2)$,

– **Case [$u_1 \neq u_2$, $v_1 \neq v_2$]:** The two edges are independent so

$$\mathbf{E}\left[\, X_i X_j \,\right] = \frac{1}{2} \cdot \frac{1}{2}.$$

– **Case [$u_1 \neq u_2$, $v_1 = v_2$]:** The two edges share a vertex. They both cross the cut iff the two different vertices are in the oposite sets, which again happens with probability 1/2

$$\mathbf{E}\left[\, X_i X_j \,\right] = \frac{1}{2} \cdot \frac{1}{2}.$$

By combining the cases, we have that

$$\mathbf{Var}\left[\, X \,\right] = |E|/2 + |E| \cdot (|E| - 1)/4 - |E|^2/4 = |E|/4.$$