Homework 4 (due Thursday, June 6th at 11:59pm)

Total points: 110 points. Maximum score: 100 points.

1. (10 pts) Show that for the r.v. $X \sim \mathcal{N}(0,1)$, we have

$$E\left[e^{tX}\right] = \mathrm{e}^{\frac{1}{2}t^2}$$

where $t \in \mathbb{R}$.

Solution:

$$\begin{aligned}
&\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}x^2\right\} \exp\{tx\} dx \\
&= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}x^2 + tx\right\} dx \\
&= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(x-t)^2\right\} \exp\left\{\frac{1}{2}t^2\right\} dx \\
&= \exp\left\{\frac{1}{2}t^2\right\} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(x-t)^2\right\} dx \\
&= \exp\left\{\frac{1}{2}t^2\right\}.
\end{aligned}$$

2. (10 pts) Let $X \sim \mathrm{Bin}(n,p)$ and $Y \sim \mathrm{Poisson}(\lambda)$. Show that

   (a) $V[X] = np(1-p)$. For what value of p is this variance maximal? Explain why this is reasonable.

   (b) $E[Y] = V(Y) = \lambda$.

Solution:

(a) To compute the variance without using the probability distribution of the binomial, we notice that since X is the number of successes in n trials, we can rewrite $X = I_1 + \ldots + I_n$, where each I_k indicates whether or not the k^{th} trial is a success. Thus, $I_k = 1$ with probability p and equals 0 otherwise.

Also, $V[I_k] = E\left[(I_k - E(I_k))^2\right] = E\left[(I_k - p)^2\right] = (1-p)^2 * P(I_k = 1) + (0-p)^2 * P(I_k = 0)$
$= (1-p)^2 p + p^2(1-p) = p(1-p)$

Now:

$$\begin{aligned}
V[X] &= V[\textstyle\sum_{k=1}^n I_k] & \\
&= \textstyle\sum_{k=1}^n V[I_k] & \text{Since all the } I_k\text{'s are independent} \\
&= nV[I_1] & \text{Since all the indicators have the same distribution} \\
&= np(1-p) & \text{Using } V[I_1] \text{ from earlier}
\end{aligned}$$

This is maximized at $p = 1/2$ since:
$\frac{d}{dp} np(1-p) = n(1-2p)$, and setting to zero and solving gives $p = 1/2$. The second derivative is $-2n$ so that this is a maximum.

This result makes sense because there is most variability in the result if the coin is fair – heads and tails are equally likely. On the other extreme, if the coin is perfectly unfair (eg $p = 0$), you will always get heads and thus there is absolutely no variation in the result ($X = 0$ always). Now, if p is close to zero, there is very little variability in the results.

(b)

$$\begin{aligned}
E(Y) &= \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} k \\
&= \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} e^{-\lambda} \\
&= \lambda \sum_{m=0}^{\infty} \frac{\lambda^m}{m!} e^{-\lambda} && \text{Reindex } m = k-1. \\
&= \lambda \underbrace{P_Y(m)}_{=1} = \lambda
\end{aligned}$$

Note that to compute $V[Y]$, we can use $V[Y] = E[Y^2] - E[Y]^2$, so we just need to compute $E[Y^2]$.

$$
\begin{aligned}
E[Y^2] &= \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} \times k^2 \\
&= \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} e^{-\lambda} k \\
&= \lambda \sum_{m=0}^{\infty} \frac{\lambda^m}{m!} e^{-\lambda} (m+1) && \text{Reindex } m = k-1 \\
&= \lambda \left(\underbrace{\left[\frac{\lambda^m}{m!} e^{-\lambda} m \right]}_{=E[Y]=\lambda} + \underbrace{\left[\frac{\lambda^m}{m!} e^{-\lambda} \times 1 \right]}_{=1} \right) \\
&= \lambda^2 + \lambda
\end{aligned}
$$

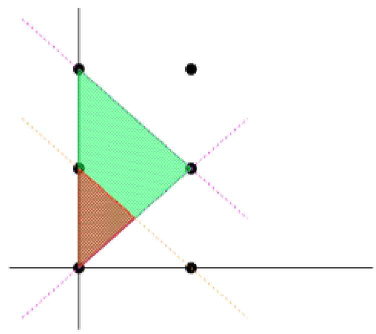
Thus, $V[Y] = (\lambda^2 + \lambda) - (\lambda)^2 = \lambda$

3. (10 pts) Consider the following function:

$$
f(x_1, x_2) = \begin{cases} 6x_1^2 x_2 \text{ if } 0 \le x_1 \le x_2 \wedge x_1 + x_2 \le 2 \\ 0 \text{ otherwise.} \end{cases}
$$

- Verify that f is a valid bivariate p.d.f.
- Suppose $(X_1, X_2) \sim f$. Are X_1 and X_2 independent random variables?
- Compute $P(Y_1 + Y_2 < 1)$
- Compute the marginal p.d.f. of X_1. What are the values of its parameters?
- Compute the marginal p.d.f. of X_2
- Compute the conditional p.d.f. of X_2 given X_1
- Compute $P(X_2 < 1.1 | X_1 = 0.6)$.

Solution:



- It is clear that $f_{X_1,X_2}(x_1,x_2) \geq 0$ for any $x_1, x_2 \in \mathbb{R}$. We just need to check that f_{X_1,X_2} integrates to 1:

$$\int_{\mathbb{R}}\int_{\mathbb{R}} f_{X_1,X_2}(x_1,x_2)\,dx_1dx_2 = \int_0^1 \int_{x_1}^{(2-x_1)} 6x_1^2 x_2\,dx_2dx_1$$
$$= 6\int_0^1 x_1^2 \int_{x_1}^{(2-x_1)} x_2\,dx_2dx_1 = 3\int_0^1 x_1^2\; x_2^2\Big|_{x_1}^{2-x_1}\,dx_1$$
$$= 3\int_0^1 x_1^2[(2-x_1)^2 - x_1^2]\,dx_1 = 12\int_0^1 x_1^2(1-x_1)\,dx_1$$
$$= 12\frac{\Gamma(3)\Gamma(2)}{\Gamma(5)} = 12\frac{2}{4!} = \frac{24}{24} = 1.$$

- X_1 and X_2 are not independent since the support of f_{X_1,X_2} (the union of the green region and the red region in the figure) is not a rectangular region.
- The region corresponding to the event $X_1 + X_2 < 1$ is the red region in the figure. We have

$$P(X_1+X_2<1) = \int_0^{\frac{1}{2}} \int_{x_1}^{1-x_1} 6x_1^2 x_2\,dx_2dx_1$$
$$= 6\int_0^{\frac{1}{2}} x_1^2 \int_{x_1}^{1-x_1} x_2\,dx_2dx_1 = 3\int_0^{\frac{1}{2}} x_1^2\; x_2^2\Big|_{x_1}^{1-x_1}\,dx_1$$
$$= 3\int_0^{\frac{1}{2}} x_1^2[(1-x_1)^2 - x_1^2]\,dx_1 = 3\int_0^{\frac{1}{2}} x_1^2(1-2x_1)\,dx_1$$
$$= x_1^3\Big|_0^{\frac{1}{2}} - \frac{3}{2}\,x_1^4\Big|_0^{\frac{1}{2}}$$
$$= \frac{1}{8} - \frac{3}{32} = \frac{1}{32}.$$

- Notice that $\mathrm{supp}(X_1) = [0,1]$. For $x_1 \in [0,1]$ we thus have

$$\begin{aligned} f_{X_1}(x_1) &= \int_{x_1}^{2-x_1} 6x_1^2 x_2 \, dx_2 = 3x_1^2 \; x_2^2 \Big|_{x_1}^{2-x_1} \\ &= 3x_1^2[(2-x_1)^2 - x_1^2] = 12x_1^2(1-x_1) \end{aligned}$$

  which we recognize as the p.d.f. of a Beta$(3,2)$ distribution.
- Notice that $\mathrm{supp}(X_2) = [0,2]$. For $0 \le x_2 < 1$ we have

$$\begin{aligned} f_{X_2}(x_2) &= \int_0^{x_2} 6x_1^2 x_2 \, dx_1 = 2x_2 \; x_1^3 \Big|_0^{x_2} \\ &= 2x_2^4. \end{aligned}$$

  For $1 \le x_2 \le 2$ we have

$$\begin{aligned} f_{X_2}(x_2) &= \int_0^{2-x_2} 6x_1^2 x_2 \, dx_1 = 2x_2 \; x_1^3 \Big|_0^{2-x_2} \\ &= 2x_2(2-x_2)^3. \end{aligned}$$

- The conditional p.d.f. of X_2 given $X_1 = x_1$ is only well-defined for $x_1 \in \mathrm{supp}(X_1) = [0,1]$. For $x_1 \in (0,1)$,

$$\begin{aligned} f_{X_2|X_1=x_1}(x_2) &= \frac{f_{X_1,X_2}(x_1,x_2)}{f_{X_1}(x_1)} = \frac{6x_1^2 x_2 \mathbb{1}_{[x_1,2-x_1]}(x_2)}{12x_1^2(1-x_1)} \\ &= \frac{x_2}{2(1-x_1)} \mathbb{1}_{[x_1,2-x_1]}(x_2). \end{aligned}$$

- We have

$$\begin{aligned} P(X_2 < 1.1 | X_1 = 0.6) &= \int_{-\infty}^{1.1} f_{X_2|X_1=0.6}(x_2) \, dx_2 \\ &= \int_{-\infty}^{1.1} \frac{x_2}{2(1-0.6)} \mathbb{1}_{[0.6,1.4]}(x_2) \\ &= \frac{1}{1.6} \; x_2^2 \Big|_{0.6}^{1.1} = \frac{1.21 - 0.36}{1.6} = 0.53125. \end{aligned}$$

4. (10 pts) Let $X_1 \sim \text{Uniform}(0,1)$ and, for $0 < x_1 \le 1$,

$$f_{X_2|X_1=x_1}(x_2) = \frac{1}{x_1} \mathbb{1}_{(0,x_1]}(x_2).$$

   - What named continuous distribution does $f_{X_2|X_1=x_1}$ seem to resemble? What are the values of its parameters?

- Compute the joint p.d.f. of (X_1, X_2)
- Compute the marginal p.d.f. of X_2.

Solution:

- It is clear that $f_{X_2|X_1=x_1}$ for $0 < x_1 \leq 1$ is a Uniform p.d.f. over the interval $(0, x_1]$.
- We have

$$\begin{aligned} f_{X_1,X_2}(x_1, x_2) = f_{X_1}(x_1) f_{X_2|X_1=x_1}(x_2) &= \mathbb{1}_{(0,1]}(x_1) \frac{1}{x_1} \mathbb{1}_{(0,x_1]}(x_2) \\ &= \frac{1}{x_1} \mathbb{1}_{\{0<x_2\leq x_1\leq 1\}}(x_1, x_2). \end{aligned}$$

- We have, for $x_2 \in (0, 1]$,

$$f_{X_2}(x_2) = \int_{x_2}^{1} \frac{1}{x_1} \, dx_1 = \log(x_1)|_{x_2}^{1} = -\log(x_2)$$

otherwise $f_{X_2}(x_2) = 0$ for $x_2 \notin (0, 1]$.

5. (10 pts) Let $(X_1, X_2) \sim f_{X_1,X_2}$ with

$$f_{X_1,X_2}(x_1, x_2) = \frac{1}{x_1} \mathbb{1}_{\{0<x_2\leq x_1\leq 1\}}(x_1, x_2).$$

Compute $E(X_1 - X_2)$.

Solution:
We have $E(X_1 - X_2) = E(X_1) - E(X_2)$. First we compute the marginal p.d.f.'s. They are

$$\begin{aligned} f_{X_1}(x_1) = \int_{\mathbb{R}} f_{X_1,X_2}(x_1, x_2) \, dx_2 &= \int_{\mathbb{R}} \frac{1}{x_1} \mathbb{1}_{\{0<x_2\leq x_1\leq 1\}}(x_1, x_2) \, dx_2 \\ &= \frac{1}{x_1} \mathbb{1}_{(0,1]}(x_1) \int_{\mathbb{R}} \mathbb{1}_{(0,x_1]}(x_2) \, dx_2 \\ &= \frac{1}{x_1} \mathbb{1}_{(0,1]}(x_1) \int_{0}^{x_1} (x_2) \, dx_2 \\ &= \mathbb{1}_{(0,1]}(x_1) \end{aligned}$$

and

$$
\begin{aligned}
f_{X_2}(x_2) &= \int_{\mathbb{R}} f_{X_1,X_2}(x_1,x_2)\,dx_1 = \int_{\mathbb{R}} \frac{1}{x_1} \mathbb{1}_{\{0<x_2\le x_1\le 1\}}(x_1,x_2)\,dx_1 \\
&= \int_{\mathbb{R}} \frac{1}{x_1} \mathbb{1}_{(0,1]}(x_2) \mathbb{1}_{[x_2,1]}(x_2)\,dx_1 \\
&= \mathbb{1}_{(0,1]}(x_2) \int_{x_2}^{1} \frac{1}{x_1}\,dx_1 = -\log(x_2)\mathbb{1}_{(0,1]}(x_2).
\end{aligned}
$$

Notice that f_{X_1} is the p.d.f. of Uniform$(0,1)$, therefore $E(X_1) = 1/2$. On the other hand

$$
\begin{aligned}
E(X_2) &= -\int_0^1 x_2 \log(x_2)\,dx_2 = -\left[\frac{x_2^2}{2}\log(x_2) - \int \frac{x_2}{2}\,dx_2\right]_0^1 \\
&= \frac{1}{4}\left[x_2^2\right]_0^1 = \frac{1}{4}.
\end{aligned}
$$

Therefore $E(X_1 - X_2) = E(X_1) - E(X_2) = 1/2 - 1/4 = 1/4$.

6. (10 pts) Let X be a random variables whose range is the set $\{0,1,2,\ldots\}$ of nonnegative integers. Show that

$$
E[X] = \sum_{i=1}^{\infty} P(X \ge i)
$$

Hint: Notice that $\sum_{i=1}^{n} P(X \ge i) = \sum_{i=1}^{\infty}\sum_{k=i}^{\infty} P(X = k)$. *Now interchange the order of summation.*

Solution:

$$
\begin{aligned}
\sum_{i=1}^{\infty} P(X \ge i) &= \sum_{i=1}^{\infty}\sum_{k=i}^{\infty} P(X = k) \\
= \sum_{1\le i\le k\le\infty} P(X = k) &= \sum_{k=1}^{\infty}\sum_{i=1}^{k} P(X = k) \\
&= \sum_{k=1}^{\infty} kP(X = k) = E(X).
\end{aligned}
$$

7. (10 pts) Let the r.v.'s X_1 and X_2 have the following joint pdf

$$
f_{X_1,X_2}(x_1,x_2) = \mathbb{1}(0 \le x_1 \le 1, 0 \le x_2 \le 1).
$$

Now let $Y = \max\{X_1, X_2\}$. Compute

(a) $P(Y \geq 3/4)$.

(b) $E[Y]$.

Solution:

(a) First, notice that the two r.v.'s are independent since the pdf can be rewritten as

$$f_{X_1,X_2}(x_1,x_2) = \mathbb{1}(0 \leq x_1 \leq 1)\mathbb{1}(0 \leq x_2 \leq 1) = f_{X_1}(x_1)f_{X_2}(x_2).$$

Then we have

$$\begin{aligned} P(Y \geq 3/4) &= 1 - P(Y < 3/4) \\ &= 1 - P(\{X_1 < 3/4\} \cap \{X_2 < 3/4\}) \\ &= 1 - P(\{X_1 < 3/4\})P(\{X_2 < 3/4\}) \\ &= 1 - \frac{3}{4}\frac{3}{4} = \frac{7}{16}. \end{aligned}$$

(b) Solution (1):

$$\begin{aligned} \int_0^1 \int_0^1 \max\{x_1,x_2\} dx_1 dx_2 &= \int_0^1 \int_0^1 [x_1 \mathbb{1}(x_1 \geq x_2) + x_2 \mathbb{1}(x_1 < x_2)] dx_1 dx_2 \\ &= \int_0^1 x_1 \int_0^1 \mathbb{1}(x_1 \geq x_2) dx_2 dx_1 + \int_0^1 x_2 \int_0^1 \mathbb{1}(x_1 < x_2) dx_1 dx_2 \\ &= \int_0^1 x_1 \int_0^{x_1} dx_2 dx_1 + \int_0^1 x_2 \int_0^{x_2} dx_1 dx_2 \\ &= \int_0^1 x_1^2 dx_1 + \int_0^1 x_2^2 dx_2 \\ &= \frac{2}{3} x^3 |_0^1 = \frac{2}{3}. \end{aligned}$$

Solution (2): notice that the pdf of Y is $F_Y(y) = y^2$. Consequently the pdf is given by $2y$. We can now compute

$$E[Y] = \int_0^1 2y \cdot y dy = 2 \int_0^1 y^2 dy = \frac{2}{3}.$$

8. (15 pts) Let the r.v.'s X_1, X_2 have joint pdf

$$f_{X_1,X_2}(x_1,x_2) = \begin{cases} cx_1x_2, \ 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1 \\ 0, \ \text{o/w} \end{cases}$$

where $c \in \mathbb{R}$. Compute

(a) the normalizing constant c;

(b) the marginal densities f_{X_1} and f_{X_2}.

(c) the conditional density $f_{X_1|X_2\geq 1/2}(x_1)$.

Now let the support of the joint pdf be $\mathbb{1}(1 \geq x_2 > x_1 \geq 0)$. Compute the same quantities as above.

Solution:

(a)

$$1 = c\int_0^1\int_0^1 x_1x_2dx_1dx_2 = c\int_0^1 x_2\int_0^1 x_1dx_1dx_2 = c\frac{1}{4} \implies c = 4.$$

(b) The r.v.'s are independent. Indeed, we can rewrite the the joint pdf as the product of the marginals

$$f_{X_1,X_2}(x_1,x_2) = 2x_1\mathbb{1}(0\leq x_1\leq 1)\cdot 2x_2\mathbb{1}(0\leq x_2\leq 1) = f_{X_1}(x_1)f_{X_2}(x_2).$$

Consequently the marginals are as above.

(c) From independence it follows that $f_{X_1|X_2\geq 1/2}(x_1) = f_{X_1}(x_1)$.

Now,

(a)

$$\begin{aligned} 1 = c\int_0^1 x_1\int_{x_1}^1 x_2dx_2dx_1 &= c\int_0^1 x_1\left.\frac{x_2^2}{2}\right|_{x_1}^1 dx_1 \\ = c\int_0^1 x_1\left(\frac{1}{2}-\frac{x_1^2}{2}\right)dx_1 &= \frac{c}{2}\left(\int_0^1 x_1dx_1 - \int_0^1 x_1^3dx_1\right) \\ = \frac{c}{2}\left(\frac{1}{2}-\frac{1}{4}\right) &= \frac{c}{2}\frac{1}{4} \implies c = 8 \end{aligned}$$

(b) For f_{X_1} we have

$$\begin{aligned} f_{X_1}(x_1) &= \int_0^1 f_{X_1,X_2}(x_1,x_2)\mathbb{1}(0\leq x_1 < x_2\leq 1)dx_2 \\ &= \int_{x_1}^1 f_{X_1,X_2}(x_1,x_2)dx_2 \\ = 8x_1\int_{x_1}^1 x_2dx_2 = 8x_1\left.\frac{x_2^2}{2}\right|_{x_1}^1 &= 8x_1\left(\frac{1}{2}-\frac{x_1^2}{2}\right) \end{aligned}$$

and for f_{X_2} we have

$$f_{X_2}(x_2) = 8x_2\frac{x_2^2}{2} = 4x_2^3.$$

(c) It's clear that first we need to obtain the marginal f_{X_2}. Therefore

$$f_{X_2}(x_2) = \int_0^{x_2} 8x_1x_2dx_1 = 8x_2\int_0^{x_2} x_1dx_1 = 8\frac{x_2^3}{2}$$

Now, use the following decomposition:

$$f_{X_1|X_2\geq 1/2}(x_1) = \frac{\int_{\max\{x_1,1/2\}}^1 8x_1x_2dx_2}{4\int_{1/2}^1 x_2^3dx_2} = \frac{8x_1\int_{\max\{x_1,1/2\}}^1 x_2dx_2}{1-\frac{1}{16}}$$

The numerator can be separated into two cases, yielding

$$f_{X_1|X_2\geq 1/2}(x_1) = \begin{cases} \frac{16}{5}x_1 \text{ if } x_1 \leq 1/2 \\ \frac{64x_1(1-x_1^2)}{15} \text{ if } x_1 > 1/2 \\ 0 \text{ o/w} \end{cases}$$

You can integrate this pdf to one to check your computations.

9. (10 pts) Let c in the following examples be the normalizing constant (different across the examples). Prove or disprove that the r.v.'s are independent:

(a) $f_{X_1,X_2}(x_1,x_2) = c\mathbb{1}(x_1 \in [0,15))\mathbb{1}(x_2 \in [0,10))$;

(b) $f_{X_1,X_2}(x_1,x_2) = ce^{-x_1-x_2-x_1x_2}\mathbb{1}(x_1 \in [0,\infty), x_2 \in [0,\infty))$;

(c) $f_{X_1,X_2}(x_1,x_2) = \frac{1}{2\sqrt{2}\pi\sigma^2}\exp\{-\frac{1}{2\sigma^2}(x_1^2+x_2^2-\sqrt{2}x_1x_2\}$

(d) $f_{X_1,X_2}(x_1,x_2) = c\frac{1}{2}\mathbb{1}(-1 < x_1 \leq 1 < x_2 < 20)$;

Solution:
For each of the distributions you need to check that $f_{X_1,X_2}(x_1,x_2) = f_{X_1}(x_1)f_{X_2}(x_2)$.

(a) Yes.

(b) No.

(c) Yes. In particular, this is a bivariate normal with covariance matrix $\sigma^2 I$ and correlation coefficient $1/\sqrt{2}$.

(d) Yes.

10. (15 pts) Let $X_1 \sim Bernoulli(p)$ and $X_2 \sim Bernoulli((X_1 + p)/2)$ for $0 \le p \le 1$. Compute $cov(X_1, X_2)$ for $p = 1/3$.

Solution:
We first need the joint pmf. This can be decomposed into

$$p_{X_1,X_2}(x_1, x_2) = p_{X_1}(x_1)p_{X_2|X_1=x_1}(x_2)$$

therefore

$$p_{X_1,X_2}(x_1, x_2) = \begin{cases} p(p+1)/2 \text{ if } x_1 = 1, x_2 = 1 \\ p(1-p)/2 \text{ if } x_1 = 1, x_2 = 0 \\ (1-p)p/2 \text{ if } x_1 = 0, x_2 = 1 \\ (1-p)(2-p)/2 \text{ if } x_1 = 0, x_2 = 0 \end{cases}$$

We also need $p_{X_2}(x_2)$. This is easy to obtain as

$$p_{X_2}(x_2) = \begin{cases} 1-p \text{ if } x_2 = 0 \\ p \text{ if } x_2 = 1 \end{cases}$$

Then $E[X_2] = 2/3$. Alternatively, you could prove this using the tower property.

$$cov(X_1, X_2) = E[X_1X_2] - E[X_1]E[X_2] = p_{X_1,X_2}(1, 1) - E[X_1]E[X_2].$$

Provided that my computations are correct, the result is

$$cov(X_1, X_2) = \frac{2}{9} - \frac{1}{9} = \frac{1}{9}.$$