PSTAT 120A, Summer 2022: Practice Problems 10: Final Review, Part I

Week 2

Conceptual Review

(a) Review the conceptual questions from the previous Discussion Worksheets!

(b) What are the different notions of conditional p.m.f.'s?

(c) What is the difference between $\mathbb{E}[X \mid Y = y]$ and $\mathbb{E}[X \mid Y]$?

(d) What is the Law of Iterated Expectation? What is the analog for variances?

1 Conditional Distributions and Expectations

Problem 1: Continuous Conditioning

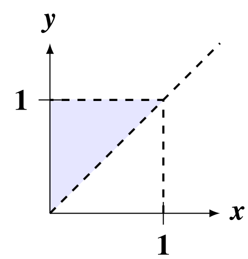
Let (X, Y) be a continuous bivariate random vector with joint p.d.f. given by

$$f_{X,Y}(x, y) = \begin{cases} c \cdot xy & \text{if } 0 < x < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

where $c > 0$ is an as-of-yet undetermined constant.

(a) Find the value of c.

Solution: We first sketch the support:



Either order of integration is fine:

$$\iint_{\mathbb{R}^2} f_{X,Y}(x, y)\, \mathrm{d}x = c \int_0^1 \int_0^y xy \, \mathrm{d}x \, \mathrm{d}y = c \int_0^1 \frac{1}{2} y^3 \, \mathrm{d}y = \frac{c}{8} \stackrel{!}{=} 1 \implies \boxed{c = 8}$$

(b) Find $f_Y(y)$, the marginal p.d.f. of Y.

Solution:

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f_{X,Y}(x, y) \, \mathrm{d}x \\ &= \int_0^y 8xy \, \mathrm{d}x = 4y^3 \implies \boxed{f_Y(y) = 4y^3 \cdot \mathbb{1}_{\{y \in [0,1]\}}} \end{aligned}$$

(c) Find $f_{X|Y}(x \mid y)$, the conditional density of X given $Y = y$.

Solution:

$$\begin{aligned} f_{X|Y}(x \mid y) &= \frac{f_{X,Y}(x, y)}{f_Y(y)} \\ &= \frac{8xy \cdot \mathbb{1}_{\{0 \le x \le y\}} \cdot \mathbb{1}_{\{y \in [0,1]\}}}{4y^3 \cdot \mathbb{1}_{\{y \in [0,1]\}}} = \boxed{\frac{2x}{y^2} \cdot \mathbb{1}_{\{0 \le x \le y\}}} \end{aligned}$$

(d) Compute $\mathbb{E}[X]$ using the Law of Iterated Expectations.

Solution:

$$\mathbb{E}[X] = \mathbb{E}\left[\mathbb{E}[X \mid Y]\right]$$
$$\mathbb{E}[X \mid Y = y] = \int_0^y \frac{2x^2}{y^2}\,\mathrm{d}x = \frac{2}{3} \cdot \frac{1}{y^2} \cdot y^3 = \frac{2}{3}Y$$
$$\mathbb{E}[X \mid Y] = \frac{2}{3}Y$$
$$\mathbb{E}[Y] = \int_0^1 4y^4\,\mathrm{d}y = \frac{4}{5}$$
$$\mathbb{E}[X] = \mathbb{E}\left[\frac{2}{3}Y\right] = \frac{2}{3}\mathbb{E}[Y] = \frac{2}{3} \cdot \frac{4}{5} = \boxed{\frac{8}{15}}$$

(e) Find $f_X(x)$, and verify your answer to part (d).

Solution:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}y = \int_x^1 8xy\,\mathrm{d}y = 4x(1-x^2) \implies \boxed{f_X(x) = 4x(1-x^2) \cdot \mathbb{1}_{\{x \in [0,1]\}}}$$
$$\mathbb{E}[X] := \int_{-\infty}^{\infty} x f_X(x)\,\mathrm{d}x = \int_0^1 4x^2(1-x^2)\,\mathrm{d}x = \frac{4}{3} - \frac{4}{5} = \frac{8}{15} \checkmark$$

Problem 2: Discrete Conditioning

Let (X, Y) be a discrete bivariate random vector with joint p.m.f. given by

$$f_{X,Y}(x, y) = \begin{cases} c \cdot xy & \text{if } x \in \{1, 2, 3\},\ y \in \{1, 2, 3\} \\ 0 & \text{otherwise} \end{cases}$$

where $c > 0$ is an as-of-yet undetermined constant.

(a) Find the value of c.

Solution:

$$\sum_x \sum_y p_{X,Y}(x, y) = c \sum_{x=1}^{3} \sum_{y=1}^{3} xy = c \cdot \frac{3 \cdot 4}{2} \cdot \frac{3 \cdot 4}{2} = 36c \overset{!}{=} 1 \implies \boxed{c = \frac{1}{36}}$$

(b) Find $p_Y(y)$, the marginal p.m.f. of Y.

Solution:

$$p_Y(y) = \sum_x p_{X,Y}(x, y) = \sum_{x=1}^{3} \frac{xy}{36} = \frac{y}{36} \cdot 6 \implies \boxed{\frac{y}{6} \cdot \mathbb{1}_{\{y \in \{1,2,3\}\}}}$$

(c) Find $p_{X|Y}(x \mid y)$, the conditional p.m.f. of X given $Y = y$.

Solution:

$$p_{X|Y}(x \mid y) = \frac{p_{X,Y}(x, y)}{p_Y(y)} = \frac{\frac{xy}{36} \cdot \mathbb{1}_{\{x \in \{1,2,3\}\}} \cdot \mathbb{1}_{\{y \in \{1,2,3\}\}}}{\frac{y}{6} \cdot \mathbb{1}_{\{y \in \{1,2,3\}\}}}$$
$$= \boxed{\frac{x}{6} \cdot \mathbb{1}_{\{x \in \{1,2,3\}\}}}$$

(d) Compute $p_X(x)$, and determine whether or not X and Y are independent. Try to make an argument using only your answer to part (c), and $p_X(x)$.

Solution:

$$p_X(x) = \sum_y p_{X,Y}(x, y) = \sum_{y=1}^{3} \frac{xy}{36} = \frac{x}{36} \cdot 6 \implies \boxed{\frac{x}{6} \cdot \mathbb{1}_{\{x \in \{1,2,3\}\}}}$$

The familiar argument goes: since $p_X(x) \cdot p_Y(y)$, we have $\boxed{X \perp Y}$. The argument involving conditional densities goes: since $p_{X|Y}(x \mid y) = p_X(x)$, $X \perp Y$.

Problem 3: Iterations!

In each of the following parts, you will be provided with the conditional distribution of $(X \mid Y)$ and the marginal distribution Y. Using the provided information, compute $\mathbb{E}[X]$ and $\mathrm{Var}(X)$.

(a) $(X \mid Y) \sim \mathrm{Bin}(Y, p); \quad Y \sim \mathrm{Pois}(\mu)$

Solution:

$$\begin{aligned}
\mathbb{E}[X \mid Y] &= Yp \\
\mathbb{E}[Y] &= \mu \\
\mathbb{E}[X] &= \mathbb{E}[\mathbb{E}[X \mid Y]] = \mathbb{E}[Yp] = p\mathbb{E}[Y] = \boxed{p\mu} \\
\mathrm{Var}(X \mid Y) &= Yp(1-p) \\
\mathrm{Var}(Y) &= \mu \\
\mathrm{Var}(X) &= \mathbb{E}[\mathrm{Var}(X \mid Y)] + \mathrm{Var}(\mathbb{E}[X \mid Y]) \\
&= \mathbb{E}[Yp(1-p)] + \mathrm{Var}(Yp) = p(1-p)\mathbb{E}[Y] + p^2\mathrm{Var}(Y) = p(1-p)\mu + p^2\mu = \boxed{p\mu}
\end{aligned}$$

(b) $(X \mid Y) \sim \mathrm{Exp}(1/Y); \quad Y \sim \mathrm{Gamma}(r, \lambda)$

Solution:

$$\begin{aligned}
\mathbb{E}[X \mid Y] &= Y \\
\mathbb{E}[Y] &= \frac{r}{\lambda} \\
\mathbb{E}[X] &= \mathbb{E}[\mathbb{E}[X \mid Y]] = \mathbb{E}[Y] = \boxed{\frac{r}{\lambda}} \\
\mathrm{Var}(X \mid Y) &= Y^2 \\
\mathrm{Var}(Y) &= \frac{r}{\lambda^2} \\
\mathrm{Var}(X) &= \mathbb{E}[\mathrm{Var}(X \mid Y)] + \mathrm{Var}(\mathbb{E}[X \mid Y]) \\
&= \mathbb{E}\left[Y^2\right] + \mathrm{Var}(Y^2) \\
&= \mathbb{E}[Y^2] + \mathbb{E}[Y^4] - (\mathbb{E}[Y^2])^2 \\
\mathbb{E}[Y^4] &= \int_0^\infty y^4 \cdot \frac{\lambda^r}{\Gamma(r)} y^{r-1} e^{-\lambda y}\, \mathrm{d}y \\
&= \frac{\lambda^r}{\Gamma(r)} \cdot \frac{\Gamma(r+4)}{\lambda^{r+4}} \cdot \int_0^\infty \frac{\lambda^{r+4}}{\Gamma(r+4)} y^{(r+4)-1} e^{-\lambda y}\, \mathrm{d}y \\
&= \frac{\Gamma(r+4)}{\lambda^4 \cdot \Gamma(r)} \\
\mathbb{E}[Y^2] &= \mathrm{Var}(Y) + (\mathbb{E}[Y])^2 = \frac{r}{\lambda^2} + \frac{r^2}{\lambda^2} = \frac{r(r+1)}{\lambda^2} \\
\mathrm{Var}(X) &= \boxed{\frac{r(r+1)}{\lambda^2} + \frac{\Gamma(r+4)}{\lambda^4 \cdot \Gamma(r)} - \frac{r^2(r+1)^2}{\lambda^4}}
\end{aligned}$$

By the way, we can simplify our $\mathbb{E}[Y^4]$ expression a bit if we so desire:

$$\frac{\Gamma(r+4)}{\lambda^4\Gamma(r)} = \frac{(r+3)(r+2)(r+1)r\Gamma(r)}{\lambda^4 \cdot \Gamma(r)} = \frac{(r+3)(r+2)(r+1)r}{\lambda^4}$$

which gives

$$\begin{aligned} \mathrm{Var}(X) &= \frac{r(r+1)}{\lambda^2} + \frac{\Gamma(r+4)}{\lambda^4 \cdot \Gamma(r)} - \frac{r^2(r+1)^2}{\lambda^4} \\ &= \frac{r(r+1)}{\lambda^2} + \frac{(r+3)(r+2)(r+1)r}{\lambda^4} - \frac{r^2(r+1)^2}{\lambda^4} \\ &= \boxed{\frac{(r+1)r(\lambda^2+4r+6)}{\lambda^4}} \end{aligned}$$

which, in addition to being a bit "cleaner" of a final expression, allows us to see that it is nonnegative for all valid values of r and λ (namely, $r > 0$ and $\lambda > 0$).

2 General Problems (includes Cond. Distn's)

Problem 4: Axiomatic Proof

Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and events $A, B \in \mathcal{F}$, prove the following identity:

$$\mathbb{P}(A \mid B^{\complement}) = 1 - \frac{\mathbb{P}(A^{\complement})}{\mathbb{P}(B^{\complement})} + \frac{\mathbb{P}(A^{\complement} \cap B)}{\mathbb{P}(B^{\complement})}$$

Solution:

$$\begin{aligned}
\mathbb{P}(A \mid B^{\complement}) &= \frac{\mathbb{P}(A \cap B^{\complement})}{\mathbb{P}(B^{\complement})} \\
&= \frac{\mathbb{P}\left[(A^{\complement} \cup B)^{\complement}\right]}{\mathbb{P}(B^{\complement})} = \frac{1 - \mathbb{P}(A^{\complement} \cup B)}{\mathbb{P}(B^{\complement})} \\
&= \frac{1 - \mathbb{P}(A^{\complement}) - \mathbb{P}(B) + \mathbb{P}(A^{\complement} \cup B)}{\mathbb{P}(B^{\complement})} \\
&= \frac{1 - \mathbb{P}(B) - \mathbb{P}(A^{\complement}) + \mathbb{P}(A^{\complement} \cup B)}{\mathbb{P}(B^{\complement})} \\
&= \frac{1 - \mathbb{P}(B)}{\mathbb{P}(B^{\complement})} - \frac{\mathbb{P}(A^{\complement})}{\mathbb{P}(B^{\complement})} + \frac{\mathbb{P}(A^{\complement} \cap B)}{\mathbb{P}(B^{\complement})} \\
&= \frac{\mathbb{P}(B^{\complement})}{\mathbb{P}(B^{\complement})} - \frac{\mathbb{P}(A^{\complement})}{\mathbb{P}(B^{\complement})} + \frac{\mathbb{P}(A^{\complement} \cap B)}{\mathbb{P}(B^{\complement})} = 1 - \frac{\mathbb{P}(A^{\complement})}{\mathbb{P}(B^{\complement})} + \frac{\mathbb{P}(A^{\complement} \cap B)}{\mathbb{P}(B^{\complement})}
\end{aligned}$$

Problem 5: Variance of Sums

Using only first principles (i.e. taking care **not** to use any previously-derived results pertaining to variance of sums of random variables), derive an expression for

$$\mathrm{Var}\left(\sum_{i=1}^{n} (-1)^i X_i \right)$$

Simplify as much as you can.

Solution:

$$\begin{aligned}
\mathrm{Var}\left(\sum_{i=1}^{n}(-1)^i X_i\right) &= \mathrm{Cov}\left(\sum_{i=1}^{n}(-1)^i X_i,\ \sum_{j=1}^{n}(-1)^j X_j\right) \\
&= \sum_{i,j}(-1)^{i+j}\mathrm{Cov}(X_i, X_j) \\
&= \sum_{i=j}(-1)^{i+j}\mathrm{Cov}(X_i, X_j) + \sum_{i \neq j}(-1)^{i+j}\mathrm{Cov}(X_i, X_j) \\
&= \sum_{i=1}^{n}\mathrm{Var}(X_i) + 2\sum_{i<j}(-1)^{i+j}\mathrm{Cov}(X_i, X_j)`
\end{aligned}$$

Problem 6: Faces of the Same Die

A fair $k-$sided die is rolled n times, where where n and k are fixed natural numbers.
Let X denote the number of faces that appear exactly three times. Find $\mathbb{E}[X]$.

Solution: We can proceed using indicators assigned to the faces of the die. Let

$$\mathbb{1}_j = \begin{cases} 1 & \text{if face } j \text{ appears exactly three times} \\ 0 & \text{otherwise} \end{cases}$$

Then

$$X = \sum_{j=1}^{k} \mathbb{1}_j$$

and so

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{j=1}^{k}\mathbb{1}_j\right] = \sum_{i=1}^{k}\mathbb{E}[\mathbb{1}_j]$$

We now turn our attention to computing $\mathbb{E}[\mathbb{1}_j]$. This amounts to computing the probability that face j (for $j = 1, 2, \cdots, k$) appears exactly twice. This can be computing using the Binomial distribution; let Y_j denote the number of times face j appears in the n rolls of the die; then $Y_j \sim \mathrm{Bin}(n, 1/k)$ and so

$$\mathbb{E}[\mathbb{1}_j] = \mathbb{P}(Y_j = 3) = \binom{n}{3}\left(\frac{1}{k}\right)^3\left(\frac{k-1}{k}\right)^{n-3}$$

Therefore,

$$\begin{aligned}\mathbb{E}[X] &= \mathbb{E}\left[\sum_{j=1}^{k} \mathbb{1}_j\right] = \sum_{j=1}^{k} \mathbb{E}[\mathbb{1}_j] \\ &= \sum_{j=1}^{k} \binom{n}{3}\left(\frac{1}{k}\right)^3\left(\frac{k-1}{k}\right)^{n-3} \\ &= k \cdot \binom{n}{3}\left(\frac{1}{k}\right)^3\left(\frac{k-1}{k}\right)^{n-3}\end{aligned}$$

Problem 7: A Useful Result

Suppose $X \sim \text{Pois}(\lambda)$ and $Y \sim \text{Pois}(\mu)$ with $X \perp Y$. Find $\mathbb{P}(X = k \mid X + Y = n)$ for natural numbers n and k with $k \leq n$, and use this to recognize the conditional distribution of $(X \mid X + Y = n)$. Be sure to include any/all relevant parameter(s)!

Hint: We know the distribution of $(X + Y)$.

Solution: We begin with the definition of conditional probability:

$$\mathbb{P}(X = k \mid X + Y = n) = \frac{\mathbb{P}(X = k,\ X + Y = n)}{\mathbb{P}(X + Y = n)}$$

Now, we know that $(X + Y) \sim \text{Pois}(\lambda + \mu)$ [this was proved on one of the Section Worksheets.] Thus, we only need to focus on the numerator. Note that

$$\mathbb{P}(X = k,\ X + Y = n) = \mathbb{P}(X = k,\ Y = n - k)$$

This is particularly useful, because we can now utilize the independence of X and Y to write

$$\begin{aligned}\mathbb{P}(X = k,\ X + Y = n) &= \mathbb{P}(X = k,\ Y = n - k) \\ &= \mathbb{P}(X = k) \cdot \mathbb{P}(Y = n - k) \\ &= e^{-\lambda}\frac{\lambda^k}{k!} \cdot e^{-\mu} \cdot \frac{\mu^{n-k}}{(n-k)!}\end{aligned}$$

Therefore, putting everything together,

$$\begin{aligned}\mathbb{P}(X = k \mid X + Y = n) &= \frac{\mathbb{P}(X = k,\ X + Y = n)}{\mathbb{P}(X + Y = n)} \\ &= \frac{\cancel{e^{-\lambda}}\frac{\lambda^k}{k!} \cdot \cancel{e^{-\mu}} \cdot \frac{\mu^{n-k}}{(n-k)!}}{\cancel{e^{-(\lambda+\mu)}} \cdot \frac{(\lambda+\mu)^n}{n!}} \\ &= \frac{n!}{k! \cdot (n-k)!} \cdot \frac{\lambda^k \mu^{n-k}}{(\lambda+\mu)^k \cdot (\lambda+\mu)^{n-k}} \\ &= \binom{n}{k}\left(\frac{\lambda}{\lambda+\mu}\right)^k\left(1 - \frac{\lambda}{\lambda+\mu}\right)^{n-k}\end{aligned}$$

Thus, we see

$$(X \mid X + Y = n) \sim \text{Bin}\left(n,\ \frac{\lambda}{\lambda + \mu}\right)$$

By the way, note that $\lambda/(\lambda + \mu) < 1$ so it is in fact a valid probability value.

Problem 8: Discrete Joint

Let (X, Y) be a discrete bivariate random vector with joint p.m.f. (probability mass function) given by

$$p_{X,Y}(x, y) = \begin{cases} c \cdot xy & \text{if } x \in \{1, 2, 3\},\ y \in \{1, 2, 3, 4\} \\ 0 & \text{otherwise} \end{cases}$$

where $c > 0$ is an as-of-yet undetermined constant.

(a) Find the value of c.

Solution: We require the joint p.m.f. to sum to unity. As such, we compute

$$\sum_{x=1}^{3} \sum_{y=1}^{4} xy = \left(\sum_{x=1}^{3} x \right) \left(\sum_{y=1}^{4} y \right) = \frac{3 \cdot 4}{2} \cdot \frac{4 \cdot 5}{2} = 60 \implies c = \frac{1}{60}$$

(b) Compute $\mathbb{P}(X = Y)$.

Solution: Note that

$$\{X = Y\} = \bigcup_k \{X = k,\ Y = k\}$$

Since the events in the union on the RHS are all disjoint, we may take the probability of both sides and invoke the third axiom of probability to see

$$\begin{aligned} \mathbb{P}(X = Y) &= \sum_k \mathbb{P}(X = k,\ Y = k) \\ &= \sum_{k=1}^{3} p_{X,Y}(k, k) = \sum_{k=1}^{3} \frac{1}{60} k^2 = \frac{1}{60} \cdot \frac{3 \cdot 4 \cdot 7}{6} = \frac{7}{30} \end{aligned}$$

Problem 9: Discrete Convolution

Let $X \sim \text{Geom}(p_1)$ and $Y \sim \text{Geom}(p_2)$ with $X \perp Y$. Derive the p.m.f. of $Z := X+Y$.
(**Note:** This will **NOT** be the Negative Binomial p.m.f., unless $p_1 = p_2$)

Solution: Let's introduce the simplifying notation $q_1 := 1 - p_1$ and $q_2 := 1 - p_2$. Then, by the Discrete Convolution formula,

$$\begin{aligned} p_Z(z) &= \sum_x p_X(x) p_Y(z - x) \\ &= \sum_x q_1^{x-1} p_1 \cdot q_2^{z-x-1} p_2 = \frac{p_1 p_2}{q_1 q_2} \cdot q_2^z \cdot \sum_x \left(\frac{q_1}{q_2} \right)^x \end{aligned}$$

Let's figure out the limits of our sum. We require both $x \in \{1, 2, \cdots\}$ and $z - x \in \{1, 2, \cdots\}$; the latter condition states $x \in \{z - 1,\ z - 2,\ \cdots\}$ meaning, combined with the first constraint, we have

$x \in \{1, \cdots, z-1\}$. Therefore:

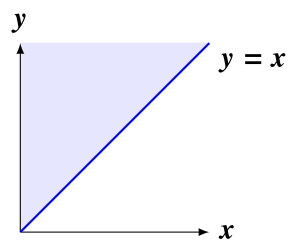
$$p_Z(z) = \frac{p_1 p_2}{q_1 q_2} \cdot q_2^z \cdot \sum_{x=1}^{z-1} \left(\frac{q_1}{q_2} \right)^x = \frac{p_1 p_2}{q_1 q_2} \cdot q_2^z \cdot \frac{\left(\frac{q_1}{q_2}\right) - \left(\frac{q_1}{q_2}\right)^z}{1 - \left(\frac{q_1}{q_2}\right)}$$

where, of course, $p_Z(z) = 0$ whenever $z \notin \{2, 3, \cdots\}$.

Problem 10: Waitin' in Line

Alex and Drew are waiting in two separate lines at *Dean* Coffee. Suppose that the time it takes for Alex to reach the counter follows an $\mathrm{Exp}(\lambda_A)$ distribution and the time it takes for Drew to reach the counter $\mathrm{Exp}(\lambda_D)$ distribution. Further suppose that the two lines move independently of each other. What is the probability that Alex reaches the counter before Drew does?

Solution: Let X denote Alex's waiting time and let Y denote Drew's waiting times. Then $X \sim \mathrm{Exp}(\lambda_A)$ and $Y \sim \mathrm{Exp}(\lambda_D)$ with $X \perp Y$. Additionally, we see $\mathbb{P}(X < Y)$, which is a double integral of $f_{X,Y}(x, y)$ over the region:



Additionally, since $X \perp Y$ we have that $f_{X,Y}(x, y) = f_X(x) \cdot f_Y(y) = \lambda_A \lambda_D e^{-\lambda_A \cdot x - \lambda_D \cdot y} \cdot \mathbb{1}_{\{x \geq 0,\ y \geq 0\}}$ meaning

$$\begin{aligned}
\mathbb{P}(X < Y) &= \int_0^\infty \int_x^\infty \lambda_A \lambda_D e^{-\lambda_A \cdot x - \lambda_D \cdot y} \, \mathrm{d}y \, \mathrm{d}x \\
&= \lambda_A \int_0^\infty e^{-\lambda_A \cdot x} \int_x^\infty \lambda_D e^{-\lambda_D \cdot y} \, \mathrm{d}y \, \mathrm{d}x \\
&= \lambda_A \int_0^\infty e^{-\lambda_A \cdot x} e^{-\lambda_D \cdot x} \, \mathrm{d}x = \lambda_A \int_0^\infty e^{-(\lambda_A + \lambda_D) x} \, \mathrm{d}x = \frac{\lambda_A}{\lambda_A + \lambda_D}
\end{aligned}$$

Note, as a quite sanity check, that this quantity is always bounded above by 1 and below by 0, meaning it is a valid probability.

Problem 11: Gamma Gamma Gamma

(a) Show that for any $r > 0$, $\Gamma(r) = (r-1)\Gamma(r-1)$

Solution: Start with

$$\Gamma(r) := \int_0^\infty x^{r-1}e^{-x}\ \mathrm{d}x$$

Integrate by parts with $u = x^{r-1}$ and $\mathrm{d}v = e^{-x}$, so $\mathrm{d}u = (r-1)x^{r-2}$ and $\mathrm{v} = -e^{-x}$, so

$$\begin{aligned}
\Gamma(r) &:= \int_0^\infty x^{r-1}e^{-x}\ \mathrm{d}x \\
&= \cancelto{0}{\left[-x^{r-1}e^{-x}\right]_{x=0}^{x=\infty}} + \int_0^\infty (r-1)x^{r-2}e^{-x}\ \mathrm{d}x \\
&= (r-1) + \int_0^\infty x^{(r-1)-1}e^{-x}\ \mathrm{d}x =: (r-1)\Gamma(r-1)
\end{aligned}$$

(b) Use part (a) to argue that $\Gamma(n) = (n-1)!$ whenever $n \in \mathbb{N}$.

Solution: If $n \in \mathbb{N}$, then

$$\begin{aligned}
\Gamma(n) &= (n-1)\Gamma(n-1) \\
&= (n-1)(n-2)\Gamma(n-2) \\
&= (n-1)(n-2)(n-3)\Gamma(n-3) \\
&\vdots \\
&= (n-1)(n-2)(n-3)\cdots 3 \times 2 \times 1 = (n-1)!
\end{aligned}$$

(c) Show that $\Gamma(1/2) = \sqrt{\pi}$.

Hint: Relate the integral to a Normal density.

Solution: Definitionally,

$$\Gamma(1/2) = \int_0^\infty x^{-1/2}e^{-x}\ \mathrm{d}x = \int_0^\infty \frac{1}{\sqrt{x}}e^{-x}\ \mathrm{d}x$$

Let $u = \sqrt{x}$ so that $\mathrm{d}u = 1/(2\sqrt{x})\ \mathrm{d}x$ and hence $\mathrm{d}x = 2u\ \mathrm{d}u$. Then:

$$\begin{aligned}
\Gamma(1/2) &= \int_0^\infty \frac{1}{\sqrt{x}}e^{-x}\ \mathrm{d}x \\
&= \int_0^\infty \frac{1}{u}e^{-u^2} \cdot 2u\ \mathrm{d}u \\
&= 2\int_0^\infty e^{-u^2}\ \mathrm{d}u = \int_{-\infty}^\infty e^{-u^2}\ \mathrm{d}u \\
&= \int_{-\infty}^\infty e^{-\frac{1}{2\cdot(1/2)}u^2}\ \mathrm{d}u
\end{aligned}$$

We recognize the integrand as the variable portion of a $\mathcal{N}(0, 1/2)$ distribution, meaning we should write

$$\begin{aligned}\Gamma(1/2) &= \int_{-\infty}^{\infty} e^{-\frac{1}{2\cdot(1/2)}u^2}\ \mathrm{d}u \\ &= \sqrt{2\pi \cdot \frac{1}{2}} \cdot \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi \cdot \frac{1}{2}}} \exp\left\{-\frac{1}{2 \cdot \frac{1}{2}} u^2\right\} \mathrm{d}u}_{1} = \sqrt{2\pi \cdot \frac{1}{2}} = \boxed{\sqrt{\pi}}\end{aligned}$$

(d) Compute $\Gamma(5/2)$.

Solution: We use parts (a) and (c) in conjunction:

$$\Gamma\left(\frac{5}{2}\right) = \frac{3}{2} \cdot \frac{1}{2} \cdot \Gamma\left(\frac{1}{2}\right) = \boxed{\frac{3\sqrt{\pi}}{4}}$$

Problem 12: It's Giving... Gamma

Prove the following identity:

Hint: Don't try to prove this directly; use probability!

$$\frac{\mathrm{d}^n}{\mathrm{d}t^n}\left(\frac{\lambda}{\lambda - t}\right)^r \bigg|_{t=0} = \frac{\Gamma(r+n)}{\Gamma(r) \cdot \lambda^n}$$

Solution: Notice that the quantity on the LHS is the n^{th} moment of a $\text{Gamma}(r, \lambda)$ distribution. Therefore, set $X \sim \text{Gamma}(r, \lambda)$; let's see if we can compute $\mathbb{E}[X^n]$ directly.

$$\begin{aligned}\mathbb{E}[X^n] &= \int_0^{\infty} x^n \cdot \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x}\ \mathrm{d}x \\ &= \frac{\lambda^r}{\Gamma(r)} \cdot \int_0^{\infty} x^{(r+n)-1} e^{-\lambda x}\ \mathrm{d}x \\ &= \frac{\lambda^r}{\Gamma(r)} \cdot \frac{\Gamma(r+n)}{\lambda^{r+n}} \cdot \int_0^{\infty} \frac{\lambda^{r+n}}{\Gamma(r+n)} x^{(r+n)-1} e^{-\lambda x}\ \mathrm{d}x \\ &= \frac{\lambda^r}{\lambda^{r+n}} \cdot \frac{\Gamma(r+n)}{\Gamma(r)} = \frac{\Gamma(r+n)}{\Gamma(r) \cdot \lambda^n}\end{aligned}$$

Therefore, we have

$$\frac{\mathrm{d}^n}{\mathrm{d}t^n}\left(\frac{\lambda}{\lambda - t}\right)^r \bigg|_{t=0} = \mathbb{E}[X^n] = \frac{\Gamma(r+n)}{\Gamma(r) \cdot \lambda^n}$$

thereby completing the proof.

Problem 13: Sums *(ASV, 9.21)*

Let $X_1, \cdots, X_{500}$ be i.i.d. random variables with expected value 2 and variance 3. The random variables $Y_1, \cdots, Y_{500}$ are independent of the X_i variables, also i.i.d., but they have expected value 2 and variance 2. Use the CLT to estimate

$$\mathbb{P}\left(\sum_{i=1}^{500} X_i > \sum_{i=1}^{500} Y_i + 50\right)$$

Solution: Let $Z_i = X_i - Y_i$. Then

$$\begin{aligned} \mathbb{E}[Z_i] &= \mathbb{E}[X_i] - \mathbb{E}[Y_i] = 2 - 2 = 0 \\ \mathrm{Var}(Z_i) &= \mathrm{Var}(X_i - Y_i) = \mathrm{Var}(X_i) + \mathrm{Var}(Y_i) = 3 + 2 = 5 \end{aligned}$$

We have

$$\mathbb{P}\left(\sum_{i=1}^{500} X_i > \sum_{i=1}^{500} Y_i + 50\right) = \mathbb{P}\left(\sum_{i=1}^{500} Z_i > 50\right)$$

Applying the central limit theorem we get

$$\begin{aligned} \mathbb{P}\left(\sum_{i=1}^{500} Z_i > 50\right) &= \mathbb{P}\left(\frac{\sum_{i=1}^{500} Z_i}{\sqrt{500 \cdot 5}} > \frac{50}{\sqrt{500 \cdot 5}}\right) \\ &\approx 1 - \Phi\left(\frac{50}{\sqrt{500 \cdot 5}}\right) = 1 - \Phi(1) = \Phi(-1) \approx 0.1587 \end{aligned}$$

Problem 14: Wald's Identity

Prove **Wald's Identity:** if $X_1, X_2, \cdots$ are i.i.d. random variables with finite mean, and N is a nonnegative integer-valued random variable independent of the X_i's (also with finite mean), then

Hint: Condition on $\{N = n\}$

$$\mathbb{E}\left[\sum_{i=1}^{N} X_i\right] = \mathbb{E}[N] \cdot \mathbb{E}[X_1]$$

Note that we cannot directly apply linearity, since the upper index of summation is random.

Solution: Following the hint, we first compute

$$\mathbb{E}\left[\sum_{i=1}^{N} X_i \,\middle|\, N = n\right] = \sum_{i=1}^{n} \mathbb{E}[X_i] = n \cdot \mathbb{E}[X_1]$$

which means

$$\mathbb{E}\left[\sum_{i=1}^{N} X_i \,\middle|\, N\right] = N \cdot \mathbb{E}[X_1]$$

Therefore, applying the Law of Iterated Expectations,

$$\begin{aligned} \mathbb{E}\left[\sum_{i=1}^{N} X_i\right] &= \mathbb{E}\left\{\mathbb{E}\left[\sum_{i=1}^{N} X_i \,\middle|\, N\right]\right\} \\ &= \mathbb{E}[N \cdot \mathbb{E}[X_1]] = \mathbb{E}[N] \cdot \mathbb{E}[X_1] \end{aligned}$$

where we have utilized the independence of N and the X_i's.

Problem 15: Let's Chalk About It *(modified from ASV 6.3)*

For each lecture, a professor chooses between white, yellow, and purple chalk, independently of previous choices. Each day she chooses white chalk with probability 0.5, yellow chalk with probability 0.4, and purple chalk with probability 0.1.

(a) Suppose we observe this professor for the next 10 days. Define appropriate random variables to count the number of times the professor chooses each color of chalk. Identify **by name** the marginal distributions, taking care to include any/all relevant parameter(s)!

Solution: Let X_W denote the number of times the professor chooses white chalk; X_Y denote the number of times the professor chooses yellow chalk, and X_P denote the number of times the professor chooses purple chalk. Then

$$\begin{aligned} X_W &\sim \text{Bin}(10, 0.5) \\ X_Y &\sim \text{Bin}(10, 0.4) \\ X_P &\sim \text{Bin}(10, 0.1) \end{aligned}$$

(b) Identify the joint distribution of the random variables you defined in part (a) **by name** (yes, it is a distribution we have encountered before). Be sure to include any/all relevant parameter(s)!

Solution: We can see that (X_W, X_Y, X_P) follows the **Multinomial Distribution**. Specifically,

$$(X_W, X_Y, X_P) \sim \text{Multinomial}(n = 10,\ r = 3,\ p_W = 0.5,\ p_Y = 0.4,\ p_P = 0.1)$$

(c) What is the probability that over the next 10 days she will choose white chalk 5 times, yellow chalk 4 times, and purple chalk 1 time?

Solution: Using part (b), we find

$$\mathbb{P}(X_W = 5, X_Y = 4, X_P = 1) = \binom{10}{5, 4, 1} \cdot \left(\frac{1}{2}\right)^5 \cdot \left(\frac{2}{5}\right)^4 \cdot \left(\frac{1}{10}\right)^1 = \frac{63}{625} \approx 0.1008$$

(d) What is the probability that over the next 10 days she will choose white chalk exactly 9 times?

Solution: Using the same notation as in part (c), we wish to compute $\mathbb{P}(X_W = 9)$. By part (a) we know that $X_W \sim \text{Bin}(10, 0.5)$ so

$$\mathbb{P}(X_W = 9) = \binom{10}{9} \cdot \left(\frac{1}{2}\right)^{10} = \frac{5}{512} \approx 0.009766$$

Problem 16: Continuous Computations

Suppose X is a random variable that has probability density function (p.d.f.) given by

$$f_X(x) = \begin{cases} ce^{-x} & \text{if } x \geq 2 \\ 0 & \text{otherwise} \end{cases}$$

where $c > 0$ is an as-of-yet undetermined constnat.

(a) Find the value of c.

Solution: $c = e^2$ (this is actually the two-parameter exponential distribution that was briefly discussed previously)

(b) Identify the distribution of $Y := X - 2$.

Solution: $Y \sim \text{Exp}(1)$

(c) Compute $\mathbb{E}[X]$

Solution: We could integrate directly, or use part (b) to see

$$\mathbb{E}[X] = \mathbb{E}[Y + 2] = \mathbb{E}[Y] + 2 = 1 + 2 = 3$$

(d) Compute $\text{Var}(X)$

Solution: Similarly as to part (c),

$$\text{Var}(X) = \text{Var}(X + 2) = \text{Var}(X) = 1$$

(e) Find $F_X(x)$, the c.d.f. of X.

Solution:

$$F_X(x) = \mathbb{P}(X \leq x) = \mathbb{P}(Y + 2 \leq x) = \mathbb{P}(Y \leq x - 2) = \begin{cases} 1 - e^{-(x-2)} & \text{if } x \geq 2 \\ 0 & \text{otherwise} \end{cases}$$

(f) Find the density of $W := X^2$.

Solution: Using the C.D.F. Method,

$$F_W(w) := \mathbb{P}(W \leq w) = \mathbb{P}(X^2 \leq w) = F_X(\sqrt{w}) - F_X(-\sqrt{w}) = F_X(\sqrt{w})$$

$$f_W(w) = \frac{1}{2\sqrt{w}} f_X(\sqrt{w}) = \frac{1}{2\sqrt{w}} e^2 e^{-\sqrt{w}} \cdot \mathbb{1}_{\{\sqrt{w} \geq 2\}} = \frac{e^2}{2} \cdot \frac{1}{\sqrt{w}} \cdot e^{-\sqrt{w}} \cdot \mathbb{1}_{\{w \geq 4\}}$$

(g) Compute $\pi_{0.67}$, the 67^{th} percentile of the distribution of X.

Solution: We seek the value $\pi_{0.67}$ such that $F_X(\pi_{0.67}) = 0.67$; i.e.

$$1 - e^{-(\pi_{0.67} - 2)} = 0.67 \implies \pi_{0.67} = 2 - \ln(0.33) \approx 3.1$$

(h) Find $M_X(t)$, the MGF of X

Solution: It is perhaps easiest to again use part (b):

$$M_X(t) = M_{Y+2}(t) = e^{2t} \cdot M_Y(t) = e^{2t} \cdot \begin{cases} \frac{1}{1-t} & \text{if } t < 1 \\ \infty & \text{otherwise} \end{cases} = \begin{cases} \frac{e^{2t}}{1-t} & \text{if } t < 1 \\ \infty & \text{otherwise} \end{cases}$$

(i) Let Y be another random variable, independent of X, that has the same p.d.f. as X. Find $f_Z(z)$, the p.d.f. of $Z := X + Y$

Solution: We use the convolution:

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z - x) \, \mathrm{d}x$$

Along the way, we will need to simplify the indicator

$$\mathbb{1}_{\{x \geq 2\}} \cdot \mathbb{1}_{\{z - x \geq 2\}} = \mathbb{1}_{\{2 \leq x \leq z-2\}}$$

meaning

$$f_Z(z) = \int_2^{z-2} e^2 e^{-x} \cdot e^2 e^{-(z-x)} \, \mathrm{d}x = e^4 (z - 4) e^{-z} \cdot \mathbb{1}_{\{z \geq 4\}}$$

(j) Suppose $\{X_i\}_{i=1}^{\infty}$ is an i.i.d. collection of random variables, following the distribution with p.d.f. given by $f_X(x)$ above. If $\bar{X}_{100}$ denotes the sample mean of 100 of these X_i's, approximate $\mathbb{P}(\bar{X}_n > 3)$.

Solution: We use the CLT to conclude that

$$\bar{X}_{100} \overset{d}{\approx} \mathcal{N}\left(3, \ \frac{1}{100}\right)$$

Therefore,

$$\mathbb{P}(\bar{X}_n > 3) = \mathbb{P}\left(\frac{\bar{X}_n - 3}{1/10} > 0\right) \approx \Phi(0) = 0.5$$

Problem 17: A Useful Result

Let $\{X_i\}$ be an i.i.d. collection of random variables with mean μ and variance σ^2.
Show that

$$\mathbb{E}\left[\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2\right] = \sigma^2$$

where $\bar{X}_n := n^{-1}\sum_{i=1}^{n} X_i$ denotes the sample mean.

Solution: First note, by Linearity,

$$\mathbb{E}\left[\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2\right] = \frac{1}{n-1}\sum_{i=1}^{n}\mathbb{E}[(X_i - \overline{X}_n^2]$$

This, we first compute

$$\begin{aligned}
\mathbb{E}[(X_i - \overline{X}_n)^2] &= \mathbb{E}[(X_i - \mu + \mu - \overline{X}_n)^2] \\
&= \mathbb{E}[(X_i - \mu)^2] + \mathbb{E}[(\overline{X}_n - \mu)^2] + 2\mathbb{E}[(X_i - \mu)(\mu - \overline{X}_n)] \\
&= \mathrm{Var}(X_i) + \mathrm{Var}(\overline{X}_n) + 2\mathbb{E}[(X_i - \mu)(\mu - \overline{X}_n)] \\
&= \sigma^2 + \frac{\sigma^2}{n} + 2\mathbb{E}[(X_i - \mu)(\mu - \overline{X}_n)]
\end{aligned}$$

Hence,

$$\begin{aligned}
\mathbb{E}\left[\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2\right] &= \frac{1}{n-1}\sum_{i=1}^{n}\mathbb{E}[(X_i - \overline{X}_n^2] \\
&= \frac{1}{n-1}\sum_{i=1}^{n}\left[\sigma^2 + \frac{\sigma^2}{n} + 2\mathbb{E}[(X_i - \mu)(\mu - \overline{X}_n)]\right] \\
&= \frac{1}{n-1}\left\{n\sigma^2 + \sigma^2 + 2\sum_{i=1}^{n}\mathbb{E}[(X_i - \mu)(\mu - \overline{X}_n)]\right\} \\
&= \frac{1}{n-1}\left\{n\sigma^2 + \sigma^2 + 2\mathbb{E}\left[\sum_{i=1}^{n}[(X_i - \mu)(\mu - \overline{X}_n)]\right]\right\} \\
&= \frac{1}{n-1}\left\{n\sigma^2 + \sigma^2 + 2\mathbb{E}\left[(\mu - \overline{X}_n)\sum_{i=1}^{n}(X_i - \mu)\right]\right\} \\
&= \frac{1}{n-1}\left\{n\sigma^2 + \sigma^2 + 2\mathbb{E}\left[(\mu - \overline{X}_n)n(\overline{X}_i - \mu)\right]\right\} \\
&= \frac{1}{n-1}\left\{n\sigma^2 + \sigma^2 - 2n\mathbb{E}\left[(\overline{X}_n - \mu)(\overline{X}_i - \mu)\right]\right\} \\
&= \frac{1}{n-1}\left\{n\sigma^2 + \sigma^2 - 2n\mathbb{E}\left[(\overline{X}_n - \mu)^2\right]\right\} \\
&= \frac{1}{n-1}\left\{n\sigma^2 + \sigma^2 - 2n\frac{\sigma^2}{n}\right\} \\
&= \frac{1}{n-1}\left[n\sigma^2 + \sigma^2 - 2\sigma^2\right] = \frac{1}{n-1}(n-1)\sigma^2 = \sigma^2
\end{aligned}$$