MTL108: Hints/ Solution to Problem Set-4

Prepared by TAs

IIT delhi

Problems

1. Define an indicator random variable and its use in computing probability. Section 4.4 in IPB.

   Solution: An indicator random variable for an event A is defined as:

   $$\mathbf{1}_A = \begin{cases} 1, & \text{if } A \text{ occurs} \\ 0, & \text{if } A \text{ does not occur} \end{cases}$$

   It indicates whether an event A is occurred or not and follows Bernoulli distribution where success means "A occurs" and failure means "A doesn't occur".

   There is a one-to-one correspondence between events and indicator random variables, and the probability of an event A is the expected value of its indicator random variable I_A:

   $$P(A) = \mathbb{E}(I_A).$$

   For any event A, we have an indicator random variable I_A. This is a one-to-one correspondence since A uniquely determines I_A and vice versa (to get from I_A back to A, we use the fact that

   $$A = \{s \in S : I_A(s) = 1\}.$$

   Since $I_A \sim \text{Bern}(p)$ with $p = P(A)$, we have

   $$\mathbb{E}(I_A) = P(A).$$

2. Let X be a hypergeometric random variable. Derive $\mathbb{E}[X]$ and $\text{Var}(X)$. Section 3.4 in IPM.

   Solution: The hypergeometric distribution models the number of successes in a sample drawn **without replacement** from a finite population.

   Suppose the population size is N and the no. of success states in the population is K. Then the no. of failures in the population is $N - K$. If we draw a sample of size n without replacement, then

   Let X be the number of successes in the sample. Then X follows a hypergeometric distribution:

   $$X \sim \text{Hypergeometric}(N, K, n).$$

   The probability that exactly k successes are observed is:

$$P(X = k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}},$$

where

$$\max(0, n - (N - K)) \leq k \leq \min(n, K).$$

The expectation and variance is:

$$\mathbb{E}[X] = n\frac{K}{N}$$

$$\text{Var}(X) = n\frac{K}{N}\left(1 - \frac{K}{N}\right)\frac{N-n}{N-1}$$

Derivation of expectation and variance:

$$\mathbb{E}[X] = \sum_k kP(X = k) = \frac{1}{\binom{N}{n}}\sum_k k\binom{K}{k}\binom{N-K}{n-k}.$$

Using the identity

$$k\binom{K}{k} = K\binom{K-1}{k-1},$$

we obtain

$$\mathbb{E}[X] = \frac{K}{\binom{N}{n}}\sum_k \binom{K-1}{k-1}\binom{N-K}{n-k}.$$

Letting $r = k - 1$ and applying Vandermonde's identity (first choosing from $K - 1$ options, then from $N - K$ is equivalent to choosing $n - 1$ from $N - 1$),

$$\sum_r \binom{K-1}{r}\binom{N-K}{n-1-r} = \binom{N-1}{n-1},$$

we get

$$\mathbb{E}[X] = \frac{K\binom{N-1}{n-1}}{\binom{N}{n}}.$$

Since

$$\frac{\binom{N-1}{n-1}}{\binom{N}{n}} = \frac{n}{N},$$

we have

$$\mathbb{E}[X] = n\frac{K}{N}.$$

We have

$$\text{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2.$$

Write X^2 as

$$X^2 = X(X - 1) + X.$$

Hence,

$$\mathbb{E}[X^2] = \mathbb{E}[X(X - 1)] + \mathbb{E}[X].$$

Now,

$$\mathbb{E}[X(X-1)] = \sum_k k(k-1)P(X=k).$$

Using the identity

$$k(k-1)\binom{K}{k} = K(K-1)\binom{K-2}{k-2},$$

we obtain

$$\mathbb{E}[X(X-1)] = \frac{K(K-1)}{\binom{N}{n}} \sum_k \binom{K-2}{k-2}\binom{N-K}{n-k}.$$

Letting $r = k-2$ and applying Vandermonde's identity again,

$$\sum_r \binom{K-2}{r}\binom{N-K}{n-2-r} = \binom{N-2}{n-2},$$

we get

$$\mathbb{E}[X(X-1)] = \frac{K(K-1)\binom{N-2}{n-2}}{\binom{N}{n}}.$$

Since

$$\frac{\binom{N-2}{n-2}}{\binom{N}{n}} = \frac{n(n-1)}{N(N-1)},$$

it follows that

$$\mathbb{E}[X(X-1)] = n(n-1)\frac{K(K-1)}{N(N-1)}.$$

Therefore,

$$\mathbb{E}[X^2] = n(n-1)\frac{K(K-1)}{N(N-1)} + n\frac{K}{N}.$$

Subtracting $(\mathbb{E}[X])^2$ and simplifying, we get

$$\mathrm{Var}(X) = n\frac{K}{N}\left(1-\frac{K}{N}\right)\frac{N-n}{N-1}.$$

3. Let $X_n \sim \text{Binomial}(n, p_n)$, $p_n \to 0$ as $n \to \infty$ and $np_n \to \lambda > 0$ as $n \to \infty$. Show that as $n \to \infty$, the distribution of X_n converges to Poisson(λ). *Hint:* Use Stirling's approximation.

Solution: Using Stirling's approximation: The binomial probability mass function is given by

$$P(X_n = k) = \binom{n}{k} p_n^k (1-p_n)^{n-k}.$$

Write the binomial coefficient using factorials:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Using Stirling's approximation:

$$m! \sim \sqrt{2\pi m}\left(\frac{m}{e}\right)^m \quad \text{as } m \to \infty.$$

Applying this to $n!$ and $(n-k)!$ (with fixed k),

$$\frac{n!}{(n-k)!} \sim \frac{\sqrt{2\pi n}\left(\frac{n}{e}\right)^n}{\sqrt{2\pi(n-k)}\left(\frac{n-k}{e}\right)^{n-k}}.$$

Simplifying,

$$\frac{n!}{(n-k)!} \sim n^k \left(\frac{n}{n-k}\right)^{n-k} e^{-k} \frac{\sqrt{n}}{\sqrt{n-k}}$$

Since k is fixed and $n \to \infty$,

$$\left(\frac{n}{n-k}\right)^{n-k} = \left(1 + \frac{k}{n-k}\right)^{n-k} \to e^k.$$

Hence,

$$\frac{n!}{(n-k)!} \sim n^k.$$

Therefore,

$$\binom{n}{k} \sim \frac{n^k}{k!}.$$

Now substitute into the pmf:

$$P(X_n = k) \sim \frac{n^k}{k!} p_n^k (1-p_n)^{n-k}.$$

Since $np_n \to \lambda$,

$$n^k p_n^k = (np_n)^k \to \lambda^k.$$

Next,

$$(1-p_n)^n \to e^{-\lambda},$$

and because $p_n \to 0$,

$$(1-p_n)^{-k} \to 1.$$

Thus,

$$(1-p_n)^{n-k} = (1-p_n)^n (1-p_n)^{-k} \to e^{-\lambda}.$$

Combining limits,

$$\lim_{n\to\infty} P(X_n = k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

This is the pmf of Poisson(λ).

Hence,

$$X_n \xrightarrow{d} \text{Poisson}(\lambda).$$

Without using Stirling's approximation:

For fixed k,

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!}.$$

Hence,

$$P(X_n = k) = \frac{n(n-1)\cdots(n-k+1)}{k!} p_n^k (1-p_n)^{n-k}.$$

Rewrite the product as

$$n(n-1)\cdots(n-k+1) = n^k \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\cdots\left(1 - \frac{k-1}{n}\right).$$

Thus,

$$P(X_n = k) = \frac{(np_n)^k}{k!}\left(1 - \frac{1}{n}\right)\cdots\left(1 - \frac{k-1}{n}\right)(1-p_n)^n(1-p_n)^{-k}.$$

Now take limits term by term.

Since $np_n \to \lambda$,

$$(np_n)^k \to \lambda^k.$$

Also,

$$\left(1 - \frac{j}{n}\right) \to 1 \quad \text{for each fixed } j,$$

so the finite product converges to 1.

Next, since $p_n \to 0$ and $np_n \to \lambda$,

$$(1-p_n)^n \to e^{-\lambda}.$$

Finally,

$$(1-p_n)^{-k} \to 1$$

because $p_n \to 0$.

Combining all limits,

$$\lim_{n\to\infty} P(X_n = k) = \frac{\lambda^k}{k!}e^{-\lambda}.$$

This is the probability mass function of Poisson(λ).

Hence,

$$X_n \xrightarrow{d} \text{Poisson}(\lambda).$$

4. Let X be a continuous random variable with PDF

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \qquad x > 0,\ \alpha > 0,\ \beta > 0.$$

Show that

$$\mathbb{E}\big(e^{Xt}\big) = (1 - t/\beta)^{-\alpha} \quad \text{for } t < \beta.$$

Solution: Given

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \qquad x > 0,$$

where $\alpha > 0$, $\beta > 0$.

We compute the moment generating function:

$$\mathbb{E}(e^{Xt}) = \int_0^\infty e^{xt} f(x)\, dx.$$

Substituting the pdf,

$$\mathbb{E}(e^{Xt}) = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{\alpha-1} e^{-\beta x} e^{xt}\, dx.$$

Combine the exponential terms:

$$= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{\alpha-1} e^{-(\beta-t)x}\, dx.$$

For convergence, we require $\beta - t > 0$, i.e., $t < \beta$.

Now use the Gamma integral identity:

$$\int_0^\infty x^{\alpha-1} e^{-cx} dx = \frac{\Gamma(\alpha)}{c^\alpha}, \qquad c > 0.$$

Here $c = \beta - t$. Thus,

$$\int_0^\infty x^{\alpha-1} e^{-(\beta-t)x} dx = \frac{\Gamma(\alpha)}{(\beta - t)^\alpha}.$$

Substituting back,

$$\mathbb{E}(e^{Xt}) = \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot \frac{\Gamma(\alpha)}{(\beta-t)^\alpha} = \frac{\beta^\alpha}{(\beta-t)^\alpha} = \left(\frac{\beta}{\beta - t}\right)^\alpha = \left(1 - \frac{t}{\beta}\right)^{-\alpha}.$$

Hence,

$$\boxed{\mathbb{E}(e^{Xt}) = (1 - t/\beta)^{-\alpha}, \quad t < \beta.}$$

5. Let X be a continuous random variable with PDF

$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \qquad 0 < x < 1,\ \alpha > 0,\ \beta > 0.$$

(a) Show that

$$\mathbb{E}[X^k] = \frac{\Gamma(\alpha+k)\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\alpha+\beta+k)}, \quad k > 0.$$

(b) Derive $\mathbb{E}[X]$ and $\mathrm{Var}(X)$.

(c) Show that if $X \sim \mathrm{Beta}(\alpha, \beta)$ and $Y = 1 - X$, then

$$Y \sim \mathrm{Beta}(\beta, \alpha).$$

Solution: (a) The expected value of X^k for a continuous random variable X with probability density function $f(x)$ is defined as:

$\mathbb{E}[X^k] = \int_{-\infty}^{\infty} x^k f(x)\, dx$

For the Beta distribution $X \sim \mathrm{Beta}(\alpha, \beta)$, the PDF is non-zero only for $0 < x < 1$:

$$\mathbb{E}[X^k] = \int_0^1 x^k \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}\, dx$$

$$= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 x^{(\alpha+k)-1}(1-x)^{\beta-1}\, dx$$

The integral term is the definition of the Beta function $B(\alpha+k, \beta)$:

$\int_0^1 x^{(\alpha+k)-1}(1-x)^{\beta-1}\, dx = B(\alpha+k, \beta) = \frac{\Gamma(\alpha+k)\Gamma(\beta)}{\Gamma(\alpha+k+\beta)}$

Substituting this back into the expression for the expectation:

$$\mathbb{E}[X^k] = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot \frac{\Gamma(\alpha+k)\Gamma(\beta)}{\Gamma(\alpha+k+\beta)}$$

$$= \frac{\Gamma(\alpha+k)\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\alpha+\beta+k)}$$

(b) Derivation of Mean and Variance

Mean ($\mathbb{E}[X]$):

Using the result from part (a) with $k = 1$:

$\mathbb{E}[X] = \frac{\Gamma(\alpha+1)\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\alpha+\beta+1)}$

Applying the recurrence property of the Gamma function, $\Gamma(z+1) = z\Gamma(z)$:

$$\mathbb{E}[X] = \frac{\alpha\Gamma(\alpha)\Gamma(\alpha+\beta)}{\Gamma(\alpha)(\alpha+\beta)\Gamma(\alpha+\beta)}$$

$$= \frac{\alpha}{\alpha+\beta}$$

Variance ($\mathrm{Var}(X)$):

First, find the second moment by setting $k = 2$:

$$\mathbb{E}[X^2] = \frac{\Gamma(\alpha+2)\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\alpha+\beta+2)}$$

$$= \frac{(\alpha+1)\alpha\Gamma(\alpha)\Gamma(\alpha+\beta)}{\Gamma(\alpha)(\alpha+\beta+1)(\alpha+\beta)\Gamma(\alpha+\beta)}$$

$$= \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)}$$

The variance is given by $\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$:

$$\begin{aligned}
\mathrm{Var}(X) &= \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)} - \left(\frac{\alpha}{\alpha+\beta}\right)^2 \\
&= \frac{\alpha(\alpha+1)(\alpha+\beta) - \alpha^2(\alpha+\beta+1)}{(\alpha+\beta)^2(\alpha+\beta+1)} \\
&= \frac{\alpha(\alpha^2+\alpha\beta+\alpha+\beta) - (\alpha^3+\alpha^2\beta+\alpha^2)}{(\alpha+\beta)^2(\alpha+\beta+1)} \\
&= \frac{\alpha^3+\alpha^2\beta+\alpha^2+\alpha\beta-\alpha^3-\alpha^2\beta-\alpha^2}{(\alpha+\beta)^2(\alpha+\beta+1)} \\
&= \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}
\end{aligned}$$

(c) Suppose $X \sim \mathrm{Beta}(\alpha, \beta)$ with pdf

$$f_X(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \qquad 0 < x < 1.$$

Let $Y = 1 - X$.

We use transformation of variables.

Since $Y = 1 - X$, we have

$$X = 1 - Y.$$

Also,

$$\left|\frac{dX}{dY}\right| = 1.$$

The support transforms as: if $0 < X < 1$, then $0 < Y < 1$.

Now,

$$f_Y(y) = f_X(1-y)\left|\frac{dX}{dY}\right|.$$

Substitute into the pdf:

$$f_Y(y) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}(1-y)^{\alpha-1}\big(1-(1-y)\big)^{\beta-1}.$$

Simplify:

$$1 - (1 - y) = y.$$

Thus,

$$f_Y(y) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}(1-y)^{\alpha-1}y^{\beta-1}.$$

Rearranging,

$$f_Y(y) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\beta)\Gamma(\alpha)}y^{\beta-1}(1-y)^{\alpha-1}, \qquad 0 < y < 1.$$

This is the pdf of a Beta(β, α) distribution.

Hence,

$$Y \sim \text{Beta}(\beta, \alpha).$$

6. (a) People are arriving at a party one at a time. While waiting for more people to arrive, they entertain themselves by comparing their birthdays. Let X be the number of people needed to obtain a birthday match, i.e., before person X arrives no two people have the same birthday, but when person X arrives there is a match. Find the PMF of X.

   Solution: Let X be the number of the person who creates the first birthday match. Assume there are $N = 365$ equally likely birthdays.

   A match cannot occur with only one person, so $X \geq 2$. By the Pigeonhole Principle, a match must occur by the $(N+1)$st person. Hence,

$$X \in \{2, 3, \ldots, N+1\}.$$

   For the first match to occur *exactly* at person x, the first $x - 1$ people must all have distinct birthdays, and the xth person must match one of them. Therefore,

$$\mathbb{P}(X = x) = \left(\frac{N}{N} \cdot \frac{N-1}{N} \cdots \frac{N-(x-2)}{N}\right)\frac{x-1}{N}.$$

   This can be written compactly as

$$\mathbb{P}(X = x) = \frac{P(N, x-1)}{N^{\,x-1}} \cdot \frac{x-1}{N}, \qquad x = 2, 3, \ldots, N+1.$$

   where $P(N, x-1)$ is permutation.

   (b) Let n be a positive integer and

$$F(x) = \frac{\lfloor x \rfloor}{n}$$

   for $0 \leq x \leq n$, $F(x) = 0$ for $x < 0$, and $F(x) = 1$ for $x > n$, where $\lfloor x \rfloor$ is the greatest integer less than or equal to x. Show that F is a CDF, and find the PMF that it corresponds to.

   Solution: Suppose

$$F(x) = \frac{\lfloor x \rfloor}{n}, \qquad 0 \leq x \leq n.$$

The function F is a valid CDF since it is non-decreasing, right-continuous, and satisfies

$$\lim_{x \to -\infty} F(x) = 0, \qquad \lim_{x \to \infty} F(x) = 1.$$

The corresponding PMF is obtained from the jumps of F. Since $\lfloor x \rfloor$ jumps by 1 at each integer k,

$$P(X = k) = F(k) - F(k^-) = \frac{k}{n} - \frac{k-1}{n} = \frac{1}{n}, \quad k = 1, 2, \ldots, n.$$

Thus, X follows a discrete uniform distribution on $\{1, 2, \ldots, n\}$.

(c) There are 100 prizes, with one worth \$1, one worth \$2, ..., and one worth 100. There are 100 boxes, each of which contains one of the prizes. You get 5 prizes by picking random boxes one at a time, *without replacement*. Find the PMF of how much your most valuable prize is worth (as a simple expression in terms of binomial coefficients).

Solution: You select 5 boxes uniformly at random from 100. Let M be the value of the most expensive prize selected.

The total number of possible selections is $\binom{100}{5}$. For $M = k$, one box must contain prize k, and the remaining 4 prizes must be chosen from the $k-1$ smaller prizes. Hence,

$$P(M = k) = \frac{\binom{k-1}{4}}{\binom{100}{5}}, \qquad k = 5, 6, \ldots, 100.$$

(d) Let F_1 and F_2 be CDFs, $0 < p < 1$, and

$$F(x) = pF_1(x) + (1-p)F_2(x) \quad \text{for all } x.$$

i. Show directly that F has the properties of a valid CDF (see Theorem 3.6.3). The distribution defined by F is called a *mixture* of the distributions defined by F_1 and F_2.

ii. Consider creating a random variable in the following way. Flip a coin with probability p of Heads. If the coin lands Heads, generate a random variable according to F_1; if the coin lands Tails, generate a random variable according to F_2. Show that the random variable obtained in this way has CDF F.

Solution: Let

$$F(x) = pF_1(x) + (1-p)F_2(x), \qquad 0 < p < 1.$$

Since F_1 and F_2 are valid CDFs, their convex combination is also a valid CDF: it has limits 0 and 1 at $\pm\infty$, is non-decreasing, and is right-continuous.

To generate such a random variable, flip a coin with $P(\text{Heads}) = p$. If Heads, sample from F_1; if Tails, sample from F_2. By the Law of Total Probability,

$$P(X \le x) = pF_1(x) + (1-p)F_2(x),$$

which matches the given CDF.

(e) An airline overbooks a flight, selling more tickets for the flight than there are seats on the plane (figuring that it is likely that some people will not show up). The plane has 100 seats, and 110 people have booked the flight. Each person will show up for the flight with probability 0.9, independently. Find the probability that there will be enough seats for everyone who shows up for the flight.

Solution: Let S be the number of passengers who show up. Each person shows up independently with probability 0.9, so

$$S \sim \text{Binomial}(110, 0.9).$$

The plane has 100 seats, so the probability that everyone fits is

$$P(S \leq 100) = \sum_{k=0}^{100} \binom{110}{k} (0.9)^k (0.1)^{110-k}.$$

Equivalently, this can be computed as $1 - P(S > 100)$.

(f) There are n people eligible to vote in a certain election. Voting requires registration. Decisions are made independently. Each of the n people will register with probability p_1. Given that a person registers, they will vote with probability p_2. Given that a person votes, they will vote for Kodos (who is one of the candidates) with probability p_3. What is the distribution of the number of votes for Kodos (give the PMF, fully simplified, or the name of the distribution, including its parameters)?

Solution: Each person independently registers, votes, and then votes for Kodos with probabilities p_1, p_2, and p_3, respectively. Thus, the probability that a given person votes for Kodos is

$$p = p_1 p_2 p_3.$$

Since there are n independent individuals, the total number of votes for Kodos, say V, follows a binomial distribution:

$$V \sim \text{Binomial}(n, p_1 p_2 p_3).$$

Its PMF is

$$P(V = k) = \binom{n}{k} (p_1 p_2 p_3)^k (1 - p_1 p_2 p_3)^{n-k}.$$

7. (a) In the Gregorian calendar, each year has either 365 days (a normal year) or 366 days (a leap year). A year is randomly chosen, with probability $\frac{3}{4}$ of being a normal year and $\frac{1}{4}$ of being a leap year. Find the mean and variance of the number of days in the chosen year.

Solution: Let X denote the number of days in a year. We are given

$$P(X = 365) = \frac{3}{4}, \qquad P(X = 366) = \frac{1}{4}.$$

Mean:

$$E[X] = 365\left(\frac{3}{4}\right) + 366\left(\frac{1}{4}\right) = \frac{1095 + 366}{4} = \frac{1461}{4} = 365.25.$$

Variance:

$$E[X^2] = 365^2\left(\frac{3}{4}\right) + 366^2\left(\frac{1}{4}\right) = 133407.75.$$

$$\text{Var}(X) = E[X^2] - (E[X])^2 = 133407.75 - (365.25)^2 = 0.1875.$$

(b) (a) A fair die is rolled. Find the expected value of the roll.

(b) Four fair dice are rolled. Find the expected total of the rolls.

Solution: (a) Expected value of one roll

For a fair six-sided die with outcomes $\{1, 2, 3, 4, 5, 6\}$,

$$E[X] = \frac{1+2+3+4+5+6}{6} = \frac{21}{6} = 3.5.$$

(b) Expected total of four rolls

Let X_1, X_2, X_3, X_4 be the outcomes of the four dice. By linearity of expectation,

$$E[X_1 + X_2 + X_3 + X_4] = E[X_1] + E[X_2] + E[X_3] + E[X_4].$$

$$E[\text{Total}] = 4 \times 3.5 = 14.$$

(c) A certain country has four regions: North, East, South, and West. The populations of these regions are 3 million, 4 million, 5 million, and 8 million, respectively. There are 4 cities in the North, 3 in the East, 2 in the South, and 1 in the West. Each person in the country lives in exactly one of these cities.

(i) What is the average size of a city in the country? (This is the arithmetic mean of the populations of the cities and is also the expected value of the population of a city chosen uniformly at random.)
Hint: Give the cities names (labels).

(ii) Show that without further information it is impossible to find the variance of the population of a city chosen uniformly at random. That is, the variance depends on how the people within each region are allocated between the cities in that region.

(iii) A region of the country is chosen uniformly at random, and then a city within that region is chosen uniformly at random. What is the expected population size of this randomly chosen city?
Hint: First find the selection probability for each city.

(iv) Explain intuitively why the answer to part (iii) is larger than the answer to part (i).

Solution: The total population is 20 million, and the total number of cities is 10.

(i) Average size of a city

$$\text{Average city size} = \frac{20}{10} = 2 \text{ million}.$$

(ii) Why the variance cannot be determined

The variance depends on the individual city populations x_i:

$$\text{Var}(X) = \frac{1}{10} \sum_{i=1}^{10} (x_i - 2)^2.$$

While regional totals are known, the distribution of population among cities within a region is unknown. Different allocations lead to different variances, even though the mean remains the same.

(iii) A region is chosen uniformly at random, and then a city within that region is chosen uniformly at random.

Since there are 4 regions, each region is selected with probability $\frac{1}{4}$.

The average city size(in million) within each region is:

$$\text{North: } \frac{3}{4}, \qquad \text{East: } \frac{4}{3}, \qquad \text{South: } \frac{5}{2}, \qquad \text{West: } 8.$$

Using the law of total expectation,

$$\mathbb{E}[X] = \frac{1}{4}\left(\frac{3}{4} + \frac{4}{3} + \frac{5}{2} + 8\right).$$

Converting to a common denominator,

$$\mathbb{E}[X] = \frac{1}{4}\left(\frac{9 + 16 + 30 + 96}{12}\right) = \frac{1}{4} \cdot \frac{151}{12} = \frac{151}{48}.$$

Hence, $\mathbb{E}[X] = \frac{151}{48} \approx 3.15$ million.

(iv) Intuition

In part (i), every city is equally likely. In part (iii), regions with fewer cities assign a higher selection probability to each city. Since the West has only one city with a large population, it is selected more often, increasing the expected value.

8. (a) Let F be the cumulative distribution function (CDF) of a continuous random variable, and let $f = F'$ be its probability density function (PDF).

   i. Show that the function g defined by

$$g(x) = 2F(x)f(x)$$

   is also a valid PDF.

   ii. Show that the function h defined by

$$h(x) = \frac{1}{2}f(-x) + \frac{1}{2}f(x)$$

   is also a valid PDF.

Solution: Let F be the CDF of a continuous random variable and $f = F'$ its PDF. To show that $g(x) = 2F(x)f(x)$ is a valid PDF, we verify the required conditions.

i. **Nonnegativity:** Since $0 \le F(x) \le 1$ and $f(x) \ge 0$ for all x,

$$g(x) = 2F(x)f(x) \ge 0.$$

ii. **Integration to 1:**

$$\int_{-\infty}^{\infty} 2F(x)f(x)\,dx.$$

Let $u = F(x)$, so $du = f(x)\,dx$. As $x \to -\infty$, $u \to 0$, and as $x \to \infty$, $u \to 1$. Thus,

$$\int_{-\infty}^{\infty} 2F(x)f(x)\,dx = \int_{0}^{1} 2u\,du = [u^2]_0^1 = 1.$$

Hence, $g(x)$ is a valid PDF.

Now consider

$$h(x) = \frac{1}{2}f(-x) + \frac{1}{2}f(x).$$

Since $f(x) \geq 0$, we have $h(x) \geq 0$. Moreover,

$$\int_{-\infty}^{\infty} h(x)\,dx = \frac{1}{2}\int_{-\infty}^{\infty} f(-x)\,dx + \frac{1}{2}\int_{-\infty}^{\infty} f(x)\,dx.$$

Using the substitution $u = -x$ in the first integral,

$$\int_{-\infty}^{\infty} f(-x)\,dx = 1.$$

Hence,

$$\int_{-\infty}^{\infty} h(x)\,dx = \frac{1}{2}(1) + \frac{1}{2}(1) = 1.$$

Thus, $h(x)$ is also a valid PDF.

(b) Let X be a continuous random variable with cumulative distribution function (CDF) F and probability density function (PDF) f.

i. Find the conditional CDF of X given $X > c$, where c is a constant such that $\mathbb{P}(X > c) \neq 0$. That is, find

$$\mathbb{P}(X \leq x \mid X > c)$$

for all x, in terms of F.

ii. Find the conditional PDF of X given $X > c$. (This is the derivative of the conditional CDF.)

iii. Check that the conditional PDF from part (b) is a valid PDF by showing directly that it is nonnegative and integrates to 1.

Solution: Let X be a continuous random variable with CDF F and PDF f, and let c be a constant such that $\mathbb{P}(X > c) \neq 0$.

i. By definition of conditional probability,

$$\mathbb{P}(X \leq x \mid X > c) = \frac{\mathbb{P}(X \leq x,\, X > c)}{\mathbb{P}(X > c)}.$$

If $x \leq c$, the events are disjoint, so the probability is 0. If $x > c$,

$$\mathbb{P}(X \leq x,\, X > c) = F(x) - F(c), \qquad \mathbb{P}(X > c) = 1 - F(c).$$

Thus, the conditional CDF is

$$\mathbb{P}(X \leq x \mid X > c) = \begin{cases} 0, & x \leq c, \\ \dfrac{F(x) - F(c)}{1 - F(c)}, & x > c. \end{cases}$$

ii. Differentiating with respect to x, the conditional PDF is

$$f_{X|X>c}(x) = \begin{cases} 0, & x \leq c, \\ \dfrac{f(x)}{1 - F(c)}, & x > c. \end{cases}$$

iii. Since $f(x) \geq 0$ and $1 - F(c) > 0$, the conditional PDF is nonnegative. Furthermore,

$$\int_{-\infty}^{\infty} f_{X|X>c}(x)\,dx = \frac{1}{1 - F(c)}\int_{c}^{\infty} f(x)\,dx = 1.$$

Hence, it is a valid PDF.

(c) A circle with a random radius $R \sim \text{Unif}(0, 1)$ is generated. Let A denote its area.

i. Find the mean and variance of A, without first finding the CDF or PDF of A.

ii. Find the CDF and PDF of A.

Solution: Let $R \sim \text{Unif}(0, 1)$ be the radius of a circle, and let $A = \pi R^2$ denote its area.

i.

$$\mathbb{E}[A] = \pi \mathbb{E}[R^2] = \pi \int_0^1 r^2 \, dr = \frac{\pi}{3}.$$

$$\mathbb{E}[A^2] = \pi^2 \mathbb{E}[R^4] = \pi^2 \int_0^1 r^4 \, dr = \frac{\pi^2}{5}.$$

$$\text{Var}(A) = \mathbb{E}[A^2] - (\mathbb{E}[A])^2 = \frac{4\pi^2}{45}.$$

ii. For $a \geq 0$,

$$\mathbb{P}(A \leq a) = \mathbb{P}\left(R \leq \sqrt{\frac{a}{\pi}}\right).$$

Thus,

$$F_A(a) = \begin{cases} 0, & a < 0, \\ \sqrt{\frac{a}{\pi}}, & 0 \leq a \leq \pi, \\ 1, & a \geq \pi. \end{cases}$$

Differentiating,

$$f_A(a) = \begin{cases} \frac{1}{2\sqrt{\pi a}}, & 0 < a < \pi, \\ 0, & \text{otherwise.} \end{cases}$$

(d) The Cauchy distribution has probability density function

$$f(x) = \frac{1}{\pi(1 + x^2)}, \qquad x \in \mathbb{R}.$$

Find the cumulative distribution function (CDF) of a random variable with the Cauchy PDF.

Hint: Recall that the derivative of the inverse tangent function $\arctan(x)$ *is*

$$\frac{d}{dx} \arctan(x) = \frac{1}{1 + x^2}.$$

Solution: The Cauchy distribution has PDF

$$f(x) = \frac{1}{\pi(1 + x^2)}, \qquad x \in \mathbb{R}.$$

The CDF is

$$F(x) = \int_{-\infty}^{x} \frac{1}{\pi(1 + t^2)} \, dt = \frac{1}{\pi} \arctan(x) + \frac{1}{2}.$$