1 Fishy Computations

Note 19

Assume for each part that the random variable can be modelled by a Poisson distribution.

(a) Suppose that on average, a fisherman catches 20 salmon per week. What is the probability that he will catch exactly 7 salmon this week?

(b) Suppose that on average, you go to Fisherman's Wharf twice a year. What is the probability that you will go at most once in 2024?

(c) Suppose that in March, on average, there are 5.7 boats that sail in Laguna Beach per day. What is the probability there will be *at least* 3 boats sailing throughout the *next two days* in Laguna?

(d) Denote $X \sim \text{Pois}(\lambda)$. Prove that

$$\mathbb{E}[Xf(X)] = \lambda\,\mathbb{E}[f(X+1)]$$

for any function f.

Solution:

(a) Let X be the number of salmon the fisherman catches per week. $X \sim \text{Poiss}(20 \text{ salmon/week})$, so

$$\mathbb{P}[X = 7 \text{ salmon/week}] = \frac{20^7}{7!}e^{-20} \approx 5.23 \cdot 10^{-4}.$$

(b) Similarly $X \sim \text{Poiss}(2)$, so

$$\mathbb{P}[X \leq 1] = \frac{2^0}{0!}e^{-2} + \frac{2^1}{1!}e^{-2} \approx 0.41.$$

(c) Let X_1 be the number of sailing boats on the next day, and X_2 be the number of sailing boats on the day after next. Now, we can model sailing boats on day i as a Poisson distribution $X_i \sim \text{Poiss}(\lambda = 5.7)$. Let Y be the number of boats that sail in the next two days. We are interested in $Y = X_1 + X_2$. We know that the sum of two independent Poisson random variables is Poisson (from Theorem 19.5 in lecture notes). Thus, we have $Y \sim \text{Poiss}(\lambda = 5.7 + 5.7 = 11.4)$.

$$\begin{aligned}
\mathbb{P}[Y \geq 3] &= 1 - \mathbb{P}[Y < 3] \\
&= 1 - \mathbb{P}[Y = 0 \cup Y = 1 \cup Y = 2] \\
&= 1 - (\mathbb{P}[Y = 0] + \mathbb{P}[Y = 1] + \mathbb{P}[Y = 2]) \\
&= 1 - \left(\frac{11.4^0}{0!}e^{-11.4} + \frac{11.4^1}{1!}e^{-11.4} + \frac{11.4^2}{2!}e^{-11.4}\right) \\
&\approx 0.999.
\end{aligned}$$

(d) We apply the Law of the Unconscious Statistician,

$$\begin{aligned}
\mathbb{E}[Xf(X)] &= \sum_{x=0}^{\infty} xf(x)\mathbb{P}[X=x] \\
&= \sum_{x=0}^{\infty} xf(x)\frac{e^{-\lambda}\lambda^x}{x!} \\
&= \sum_{x=1}^{\infty} xf(x)\frac{e^{-\lambda}\lambda^x}{x!} \\
&= \lambda \sum_{x=1}^{\infty} f(x)\frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!} \\
&= \lambda \sum_{x=0}^{\infty} f(x+1)\frac{e^{-\lambda}\lambda^x}{x!} \\
&= \lambda\,\mathbb{E}[f(X+1)]
\end{aligned}$$

as desired.

2 Such High Expectations

Note 19 Suppose X and Y are independently drawn from a Geometric distribution with parameter p.

(a) Compute $\mathbb{E}[\min(X,Y)]$.

(b) Compute $\mathbb{E}[\max(X,Y)]$.

Solution:

(a) By independence,

$$\mathbb{P}[\min(X,Y) \geq t] = \mathbb{P}[X \geq t]\mathbb{P}[Y \geq t] = (1-p)^{2(t-1)}.$$

By Tail Sum,

$$\mathbb{E}[\min(X,Y)] = \sum_{t=1}^{\infty} \mathbb{P}[\min(X,Y) \geq t] = \sum_{t=1}^{\infty} (1-p)^{2(t-1)} = \frac{1}{1-(1-p)^2}.$$

Alternate Solution: We can see that $\min(X,Y)$ is a geometric distribution by looking at the tail probability from earlier. In particular, we have that $\min(X,Y) \sim \text{Geom}(1-(1-p)^2)$. This means that

$$\mathbb{E}[\min(X,Y)] = \frac{1}{1-(1-p)^2},$$

from the expectation of a geometric distribution.

(b) We see that

$$\begin{aligned}\mathbb{P}[\max(X,Y) \ge t] = 1 - \mathbb{P}[\max(X,Y) < t] &= 1 - \mathbb{P}[X < t]\mathbb{P}[Y < t] \\ &= 1 - (1 - \mathbb{P}[X \ge t])(1 - \mathbb{P}[Y \ge t]) \\ &= 1 - \left(1 - (1-p)^{t-1}\right)\left(1 - (1-p)^{t-1}\right) \\ &= 1 - \left(1 - 2(1-p)^{t-1} + (1-p)^{2(t-1)}\right) \\ &= 2(1-p)^{t-1} - (1-p)^{2(t-1)}.\end{aligned}$$

Using the result from part (a),

$$\begin{aligned}\mathbb{E}[\max(X,Y)] &= \sum_{t=1}^{\infty} \mathbb{P}[\max(X,Y) \ge t] \\ &= \sum_{t=1}^{\infty} 2(1-p)^{t-1} - (1-p)^{2(t-1)} \\ &= \sum_{t=1}^{\infty} 2(1-p)^{t-1} - \sum_{t=1}^{\infty} (1-p)^{2(t-1)} \\ &= \frac{2}{p} - \frac{1}{1-(1-p)^2}.\end{aligned}$$

Alternate Solution: An extremely elegant one-liner with linearity:

$$\mathbb{E}[\max(X,Y)] = \mathbb{E}[X + Y - \min(X,Y)] = \mathbb{E}[X] + \mathbb{E}[Y] - \mathbb{E}[\min(X,Y)] = \frac{2}{p} - \frac{1}{1-(1-p)^2}.$$

3 Diversify Your Hand

Note 15
Note 16

You are dealt 5 cards from a standard 52 card deck. Let X be the number of distinct values in your hand. For instance, the hand (A, A, A, 2, 3) has 3 distinct values.

(a) Calculate $\mathbb{E}[X]$. (Hint: Consider indicator variables X_i representing whether i appears in the hand.)

(b) Calculate $\mathrm{Var}(X)$.

Solution:

(a) Let X_i be the indicator of the ith value appearing in your hand. Then, $X = X_1 + X_2 + \ldots + X_{13}$. (Here we let 13 correspond to K, 12 correspond to Q, and 11 correspond to J.) By linearity of expectation, $\mathbb{E}[X] = \sum_{i=1}^{13} \mathbb{E}[X_i]$.

We can calculate $\mathbb{P}[X_i = 1]$ by taking the complement, $1 - \mathbb{P}[X_i = 0]$, or 1 minus the probability that the card does not appear in your hand. This is $1 - \frac{\binom{48}{5}}{\binom{52}{5}}$.

Then, $\mathbb{E}[X] = 13\mathbb{P}[X_1 = 1] = 13\left(1 - \frac{\binom{48}{5}}{\binom{52}{5}}\right)$.

(b) To calculate variance, since the indicators are not independent, we have to use the formula $\mathbb{E}[X^2] = \sum_{i=j}\mathbb{E}[X_i^2] + \sum_{i\neq j}\mathbb{E}[X_iX_j]$.

First, we have

$$\sum_{i=j}\mathbb{E}[X_i^2] = \sum_{i=j}\mathbb{E}[X_i] = 13\left(1-\frac{\binom{48}{5}}{\binom{52}{5}}\right).$$

Next, we tackle $\sum_{i\neq j}\mathbb{E}[X_iX_j]$. Note that $\mathbb{E}[X_iX_j] = \mathbb{P}[X_iX_j = 1]$, as X_iX_j is either 0 or 1.

To calculate $\mathbb{P}[X_iX_j = 1]$ (the probability we have both cards in our hand), we note that $\mathbb{P}[X_iX_j = 1] = 1-\mathbb{P}[X_i = 0]-\mathbb{P}[X_j = 0]+\mathbb{P}[X_i = 0, X_j = 0]$. Then

$$\begin{aligned}
\sum_{i\neq j}\mathbb{E}[X_iX_j] &= 13\cdot 12\mathbb{P}[X_iX_j = 1] \\
&= 13\cdot 12(1-\mathbb{P}[X_i = 0]-\mathbb{P}[X_j = 0]+\mathbb{P}[X_i = 0, X_j = 0]) \\
&= 156\left(1-2\frac{\binom{48}{5}}{\binom{52}{5}}+\frac{\binom{44}{5}}{\binom{52}{5}}\right)
\end{aligned}$$

Putting it all together, we have

$$\begin{aligned}
\mathrm{Var}(X) &= \mathbb{E}[X^2]-\mathbb{E}[X]^2 \\
&= 13\left(1-\frac{\binom{48}{5}}{\binom{52}{5}}\right)+156\left(1-2\frac{\binom{48}{5}}{\binom{52}{5}}+\frac{\binom{44}{5}}{\binom{52}{5}}\right)-\left(13\left(1-\frac{\binom{48}{5}}{\binom{52}{5}}\right)\right)^2.
\end{aligned}$$

4 Swaps and Cycles

Note 15

We'll say that a permutation $\pi = (\pi(1),\ldots,\pi(n))$ contains a *swap* if there exist $i, j \in \{1,\ldots,n\}$ so that $\pi(i) = j$ and $\pi(j) = i$, where $i \neq j$.

(a) What is the expected number of swaps in a random permutation?

(b) In the same spirit as above, we'll say that π contains a *k-cycle* if there exist $i_1,\ldots,i_k \in \{1,\ldots,n\}$ with $\pi(i_1) = i_2, \pi(i_2) = i_3,\ldots,\pi(i_k) = i_1$. Compute the expectation of the number of k-cycles.

Solution:

(a) As a warm-up, let's compute the probability that 1 and 2 are swapped. There are $n!$ possible permutations, and $(n-2)!$ of them have $\pi(1) = 2$ and $\pi(2) = 1$. This means

$$\mathbb{P}[(1,2)\text{ are a swap}] = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}.$$

There was nothing special about 1 and 2 in this calculation, so for any $\{i,j\} \subset \{1,\ldots,n\}$, the probability that i and j are swapped is the same as above. Let's write $I_{i,j}$ for the indicator that i and j are swapped, and N for the total number of swaps, so that

$$\mathbb{E}[N] = \mathbb{E}\left[\sum_{\{i,j\}\subset\{1,\ldots,n\}} I_{i,j}\right] = \sum_{\{i,j\}\subset\{1,\ldots n\}} \mathbb{P}[(i,j) \text{ are swapped}] = \frac{1}{n(n-1)}\binom{n}{2} = \frac{1}{2}.$$

(b) The idea here is quite similar to the above, so we'll be a little less verbose in the exposition. However, as a first aside we need the notion of a *cyclic ordering* of k elements from a set $\{1,\ldots n\}$. We mean by this a labelling of the k beads of a necklace with elements of the set, where we say that labellings of the beads are the same if we can move them along the string to turn one into the other. For example, $(1,2,3,4)$ and $(1,2,4,3)$ are different cyclic orderings, but $(1,2,3,4)$ and $(2,3,4,1)$ are the same. There are

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

possible cyclic orderings of length k from a set with n elements, since if we first count all subsets of size k, and then all permutations of each of those subsets, we have overcounted by a factor of k.

Now, let N be a random variable counting the number of k-cycles, and for each cyclic ordering $(i_1,\ldots,i_k)$ of k elements of $\{1,\ldots,n\}$, let $I_{(i_1,\ldots,i_k)}$ be the indicator that $\pi(i_1) = i_2, \pi(i_2) = i_3, \ldots, \pi(i_k) = i_1$. There are $(n-k)!$ permutations in which $(i_1,\ldots,i_k)$ form an k-cycle (since we are free to do whatever we want to the remaining $(n-k)$ elements of $\{1,\ldots,n\}$), so the probability that $(i_1,\ldots,i_k)$ are such a cycle is $\frac{(n-k)!}{n!}$, and

$$\mathbb{E}[N] = \mathbb{E}\left[\sum_{(i_1,\ldots,i_k)\text{ cyclic ordering}} I_{(i_1,\ldots,i_k)}\right] = \frac{n!}{(n-k)!}\frac{1}{k}\frac{(n-k)!}{n!} = \frac{1}{k}.$$

5 Double-Check Your Intuition Again

Note 16

(a) You roll a fair six-sided die and record the result X. You roll the die again and record the result Y.

(i) What is $\operatorname{cov}(X+Y, X-Y)$?

(ii) Prove that $X+Y$ and $X-Y$ are not independent.

For each of the problems below, if you think the answer is "yes" then provide a proof. If you think the answer is "no", then provide a counterexample.

(b) If X is a random variable and $\operatorname{Var}(X) = 0$, then must X be a constant?

(c) If X is a random variable and c is a constant, then is $\operatorname{Var}(cX) = c\operatorname{Var}(X)$?

(d) If A and B are random variables with nonzero standard deviations and $\text{Corr}(A,B) = 0$, then are A and B independent?

(e) If X and Y are not necessarily independent random variables, but $\text{Corr}(X,Y) = 0$, and X and Y have nonzero standard deviations, then is $\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y)$?

The two subparts below are **optional** and will not be graded but are recommended for practice.

(f) If X and Y are random variables then is $\mathbb{E}[\max(X,Y)\min(X,Y)] = \mathbb{E}[XY]$?

(g) If X and Y are independent random variables with nonzero standard deviations, then is

$$\text{Corr}(\max(X,Y),\min(X,Y)) = \text{Corr}(X,Y)?$$

Solution:

(a) (i) Using bilinearity of covariance, we have

$$\begin{aligned}\text{cov}(X+Y,X-Y) &= \text{cov}(X,X) + \text{cov}(X,Y) - \text{cov}(Y,X) - \text{cov}(Y,Y)\\ &= \text{cov}(X,X) - \text{cov}(Y,Y),\\ &= 0\end{aligned}$$

where we use that $\text{cov}(X,Y) = \text{cov}(Y,X)$ to get the second equality.

(ii) Observe that $\mathbb{P}[X+Y=7, X-Y=0] = 0$ because if $X-Y=0$, then the sum of our two dice rolls must be even. However, both $\mathbb{P}[X+Y=7]$ and $\mathbb{P}[X-Y=0]$ are nonzero, so $\mathbb{P}[X+Y=7, X-Y=0] \neq \mathbb{P}[X+Y=7] \cdot \mathbb{P}[X-Y=0]$.

(b) Yes. If we write $\mu = \mathbb{E}[X]$, then $0 = \text{Var}(X) = \mathbb{E}[(X-\mu)^2]$ so $(X-\mu)^2$ must be identically 0 since perfect squares are non-negative. Thus $X = \mu$.

(c) No. We have $\text{Var}(cX) = \mathbb{E}[(cX - \mathbb{E}[cX])^2] = c^2\,\mathbb{E}[(X-\mathbb{E}[X])^2] = c^2\,\text{Var}(X)$ so if $\text{Var}(X) \neq 0$ and $c \neq 0$ or $c \neq 1$ then $\text{Var}(cX) \neq c\,\text{Var}(X)$. This does prove that $\sigma(cX) = c\sigma(X)$ though.

(d) No. Let $A = X+Y$ and $B = X-Y$ from part (a). Since A and B are not constants then part (b) says they must have nonzero variances which means they also have nonzero standard deviations. Part (a) says that their covariance is 0 which means they are uncorrelated, and that they are not independent.

Recall from lecture that the converse is true though.

(e) Yes. If $\text{Corr}(X,Y) = 0$, then $\text{cov}(X,Y) = 0$. We have $\text{Var}(X+Y) = \text{cov}(X+Y,X+Y) = \text{Var}(X) + \text{Var}(Y) + 2\,\text{cov}(X,Y) = \text{Var}(X) + \text{Var}(Y)$.

(f) Yes. For any values x,y we have $\max(x,y)\min(x,y) = xy$. Thus, $\mathbb{E}[\max(X,Y)\min(X,Y)] = \mathbb{E}[XY]$.

(g) No. You may be tempted to think that because $(\max(x,y), \min(x,y))$ is either (x,y) or (y,x), then $\mathrm{Corr}(\max(X,Y), \min(X,Y)) = \mathrm{Corr}(X,Y)$ because $\mathrm{Corr}(X,Y) = \mathrm{Corr}(Y,X)$. That reasoning is flawed because $(\max(X,Y), \min(X,Y))$ is not always equal to (X,Y) or always equal to (Y,X) and the inconsistency affects the correlation. It is possible for X and Y to be independent while $\max(X,Y)$ and $\min(X,Y)$ are not.

For a concrete example, suppose X is either 0 or 1 with probability $1/2$ each and Y is independently drawn from the same distribution. Then $\mathrm{Corr}(X,Y) = 0$ because X and Y are independent. Even though X never gives information about Y, if you know $\max(X,Y) = 0$ then you know for sure $\min(X,Y) = 0$.

More formally, $\max(X,Y) = 1$ with probability $3/4$ and 0 with probability $1/4$, and $\min(X,Y) = 1$ with probability $1/4$ and 0 with probability $3/4$. This means

$$\mathbb{E}[\max(X,Y)] = 1 \cdot \frac{3}{4} + 0 \cdot \frac{1}{4} = \frac{3}{4}$$

and

$$\mathbb{E}[\min(X,Y)] = 1 \cdot \frac{1}{4} + 0 \cdot \frac{3}{4} = \frac{1}{4}.$$

Thus,

$$\begin{aligned} \mathrm{cov}(\max(X,Y), \min(X,Y)) &= \mathbb{E}[\max(X,Y)\min(X,Y)] - \frac{3}{16} \\ &= \frac{1}{4} - \frac{3}{16} = \frac{1}{16} \neq 0 \end{aligned}$$

We conclude that $\mathrm{Corr}(\max(X,Y), \min(X,Y)) \neq 0 = \mathrm{Corr}(X,Y)$.