

[Homework] Conditional variance

1. Customers depart from a bookstore according to a Poisson process with rate $\lambda = 10$ per hour. Each customer buys a book with probability $p = 0.3$, independent of everything else. If a customer buys a book, she spend \$10 on average with the standard deviation \$3 independently.

- (a) Find the distribution of the time until the first sale of a book.

Sol. This is the time until the first customer departure in the split Poisson process. It is therefore exponentially distributed with parameter $p\lambda = 3$.

- (b) Find the probability that no books are sold during a particular hour.

Sol. This is the probability of no customers in the split Poisson process during an hour, and using the result of part (a), $e^{-3} \cdot \frac{3^0}{0!}$

- (c) Find the expected number of customers who buy a book during a particular hour.

Sol. This is the expected number of customers in the split Poisson process during an hour,

$$EN = p\lambda = 3$$

- (d) Find the expectation and the standard deviation of the sales during a particular day of 12 hours business opening.

Sol. Let X_i : amount of money that customer spend at bookstore per hour and Y_i : total amount of money that customer spend at ith time.

$$EX_i = EX = 10, VarX_i = VarX = 9$$

$$EX_iX_j = EX_iEX_j = (EX)^2 = 100 \text{ (for } i \neq j \text{)}$$

$$Y_k = X_1 + \cdots X_N$$

$$\begin{aligned} EY_k &= \sum_{i=1}^{\infty} P(N = i)E[Y_k|N = i] \\ &= \sum_{i=1}^{\infty} P(N = i)i \cdot EX \\ &= EX \cdot \sum_{i=1}^{\infty} P(N = i)i \\ &= EX \cdot EN = 10 \cdot 3 \end{aligned}$$

$$\begin{aligned}
EY_k^2 &= \sum_{i=1}^{\infty} P(N=i)E[Y_k^2|N=i] \\
&= \text{Var}X \cdot EN + (EX)^2 \cdot EN^2
\end{aligned}$$

(Why? Explain!)

$$\begin{aligned}
\text{Var}Y_k &= EY_k^2 - (EY_k)^2 \\
&= \text{Var}X \cdot EN + (EX)^2 \cdot \{EN^2 - (EN)^2\} \\
&= \text{Var}X \cdot EN + (EX)^2 \cdot \text{Var}N
\end{aligned}$$

Let Y : the sales during a particular day of 12 hours business opening.
 $Y = Y_1 + Y_2 + \cdots + Y_{11} + Y_{12}$

$$\begin{aligned}
EY &= E\left(\sum_{k=1}^{12} Y_k\right) \\
&= 12E(Y_k)
\end{aligned}$$

$$\begin{aligned}
\sigma_Y &= \sigma_{\sum_{k=1}^{12} Y_k} \\
&= 12\sigma_{Y_k}
\end{aligned}$$

2. Each morning, as you pull out of your driveway, you would like to make a U-turn rather than drive around the block. Unfortunately U-turn are illegal in your neighborhood, and police cars drive by according to a Poisson process with rate λ . You decide to make a U-turn once you see that the road has been clear of police cars for τ time units. Let N be the number of police cars you see before you make the U-turn.

(a) Find $E[N]$.

Sol. Let T_n : the time difference between the police car n and $n+1$ arrival

$$P(N=0) = P(T_1 > \tau)$$

$$P(N=1) = P(T_1 \leq \tau)P(T_2 > \tau)$$

$$P(N=k) = P(T_1 \leq \tau)P(T_2 \leq \tau) \cdots P(T_k \leq \tau)P(T_{k+1} > \tau) = q^k p$$

Since $T_n \sim \text{Exp}(\lambda)$,

$$p = P(T_n > \tau) = \int_{\tau}^{\infty} \lambda e^{-\lambda x} dx = e^{-\tau\lambda}$$

Since $N + 1 \sim \text{Geo}(p)$.

$$EN = \frac{1}{p} - 1 = e^{\lambda\tau} - 1$$

$$\text{Var}N = \frac{q}{p^2} = (1 - e^{-\lambda\tau})e^{2\lambda\tau}$$

- (b) Find the conditional expectation of the time elapsed between police cars $n - 1$ and n , given that $N \geq n$.

Sol.

$$E[T_n | N \geq n] = E[T_n | T_n \leq \tau]$$

(because of the memorylessness of the exponential PDF)

M1.

$$E(T_n | T_n > \tau) = \tau + \frac{1}{\lambda}$$

$$ET_n = E(T_n | T_n \leq \tau)P(T_n \leq \tau) + E(T_n | T_n > \tau)P(T_n > \tau)$$

$$\frac{1}{\lambda} = E(T_n | T_n \leq \tau)(1 - e^{-\lambda\tau}) + \left(\tau + \frac{1}{\lambda}\right) \cdot e^{-\lambda\tau}$$

So,

$$E(T_n | T_n \leq \tau) = \frac{\frac{1}{\lambda} - \left(\tau + \frac{1}{\lambda}\right) \cdot e^{-\lambda\tau}}{1 - e^{-\lambda\tau}}$$

M2.

$$E[T_n | N \geq n] = E[T_n | T_n \leq \tau] = \frac{\int_0^\tau s \lambda e^{-\lambda s} ds}{\int_0^\tau \lambda e^{-\lambda s} ds}$$

- (c) Find the expectation of your waiting time until you make the U-turn.

Sol. Let T be the waiting time until you make the U-turn.

$$ET_n = E(T_n | T_n \leq \tau)P(T_n \leq \tau) + E(T_n | T_n > \tau)P(T_n > \tau)$$

$$\frac{1}{\lambda} = E(T_n | T_n \leq \tau)(1 - e^{-\lambda\tau}) + \left(\tau + \frac{1}{\lambda}\right) \cdot e^{-\lambda\tau}$$

So,

$$E(T_n | T_n \leq \tau) = \frac{\frac{1}{\lambda} - \left(\tau + \frac{1}{\lambda}\right) \cdot e^{-\lambda\tau}}{1 - e^{-\lambda\tau}}$$

Let T : total waiting time until you make the U-turn

$$T = T_1 + T_2 + \cdots + T_N + \tau = \sum_{k=1}^N T_k + \tau$$

$$\begin{aligned} E(T) &= \sum_{n=0}^{\infty} P(N = n) E(T_1 + \cdots + T_N | N = n) + \tau \\ &= \sum_{n=0}^{\infty} P(N = n) \sum_{i=1}^n E(T_i | T_1 \leq \tau, \dots, T_n \leq \tau, T_{n+1} > \tau) + \tau \\ &= \sum_{n=0}^{\infty} P(N = n) \sum_{i=1}^n E(T_i | T_i \leq \tau) + \tau \\ &= \sum_{n=0}^{\infty} P(N = n) \cdot n \cdot E(T_i | T_i \leq \tau) + \tau \\ &= E(N) \cdot E(T_i | T_i \leq \tau) + \tau \end{aligned}$$

(d) Find the variance of your waiting time until you make the U-turn.

Sol.

$$\text{Var}T = \text{Var}[E(T|N)] + E[\text{Var}(T|N)]$$

3. The joint PDF of X and Y is given by

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

Compute $E[X^2 | Y = y]$.

Sol1.

$$\begin{aligned} f_{X|Y}(x|y) &= \frac{f(x, y)}{f_Y(y)} \\ &= \frac{1}{y} e^{-\frac{x}{y}} \end{aligned}$$

$$\begin{aligned} E[X^2 | Y = y] &= \int_0^{\infty} x^2 f_{X|Y}(x|y) dx \\ &= \int_0^{\infty} x^2 \cdot \frac{1}{y} e^{-\frac{x}{y}} dx \\ &= \lim_{t \rightarrow \infty} (-x^2 - 2xy - 2y^2) e^{-\frac{x}{y}} \Big|_{x=0}^{x=t} \\ &= 2y^2 \end{aligned}$$

Sol2.

$$E[X^2|Y = y] = \text{Var}(X|Y = y) + (E[X|Y = y])^2 = y^2 + y^2 = 2y^2$$

4. Let Y denote a uniform random variable on $[0, 1]$, and suppose that, conditional on $Y = p$, the random variable X_n has a binomial distribution with parameters n and p . What is the mean and variance of X_n ?

Sol.

$$\begin{aligned} P(X_n = i) &= \int_0^1 P(X_n = i|U = p) f_U(p) dp \\ &= \int_0^1 \binom{n}{i} p^i (1-p)^{n-i} \cdot 1 dp \\ &= \binom{n}{i} \int_0^1 p^i (1-p)^{n-i} \cdot 1 dp \\ &= \binom{n}{i} \frac{\Gamma(i+1)\Gamma(n-i+1)}{\Gamma(i+1+n-i+1)} \\ &= \frac{1}{n+1} \end{aligned}$$

$$EX_n = E(E(X_n|Y = p)) = E(np) = n \cdot Ep = n \cdot \frac{1}{2}$$

$$\begin{aligned} \text{Var} X_n &= \text{Var}(E(X_n|p)) + E(\text{Var}(X_n|p)) \\ &= \text{Var}(np) + E(np(1-p)) \\ &= n^2 \text{Var}(p) + n(Ep - Ep^2) \\ &= n^2 \text{Var}(p) + n(Ep - (\text{Var} p + (Ep)^2)) \\ &= n^2 \cdot \frac{1}{12} + n \cdot \left(\frac{1}{2} - \left(\frac{1}{12} + \left(\frac{1}{2} \right)^2 \right) \right) \\ &= \frac{n^2}{12} + \frac{n}{6} \end{aligned}$$

5. Ten hunters are waiting for ducks to fly by. When a flock of ducks flies overhead, the hunters fire at the same time, but each chooses his target at random, independently of the others. If each hunter independently hits his target with probability 0.6, compute the expected number of ducks that are hit, where the number of ducks in a flock is a Poisson random variable with mean 6. How about the variance?

Sol. Let N be the number of ducks in a flock, which is $Po(\lambda)$, $\lambda = 6$. For

each duck i let X_i be the indicator of the event A_i that duck i is hit; $X_i = 1_{A_i}$. Consider the number Y of ducks that are hit;

$$Y = \sum_{i=1}^N X_i.$$

$$\begin{aligned} E(1 - X_i|N) &= E(1_{A_i^c}|N) = P(A_i^c|N) \\ &= \sum_{k=0}^{10} P(H_i = k|N)P(A_i^c|N, H_i = k) \\ &= \sum_{k=0}^{10} \binom{10}{k} \left(\frac{1}{N}\right)^k \left(\frac{N-1}{N}\right)^{10-k} (0.4)^k \\ &= \sum_{k=0}^{10} \binom{10}{k} \left(\frac{0.4}{N}\right)^k \left(\frac{N-1}{N}\right)^{10-k} \\ &= \left(\frac{0.4}{N} + \frac{N-1}{N}\right)^{10} \end{aligned}$$

$$\begin{aligned} E(Y|N) &= \sum_{i=1}^N E(X_i|N) = \sum_{i=1}^N \left[1 - \left(\frac{0.4}{N} + \frac{N-1}{N}\right)^{10}\right] \\ &= N \left[1 - \left(\frac{N-0.6}{N}\right)^{10}\right] \end{aligned} \tag{0.1}$$

with $\lambda = 6$

$$EY = \sum_{N=1}^{\infty} N \left[1 - \left(\frac{N-0.6}{N}\right)^{10}\right] \frac{\lambda^N}{N!} e^{-\lambda}$$

$$E(Y^2|N) = NP(A_1) + N(N-1)P(A_1A_2) \tag{0.2}$$

6. The number of people who enter an elevator on the ground floor is a Poisson random variable with mean 10. If there are N floors above the ground floor, and if each person is equally likely to get off at any one of the N floors, independently of where the others get off, compute the expected number of stops that the elevator

will make before discharging all of its passengers.

Sol. Let

$$I_i = \begin{cases} 1 & \text{elevator stops at floor } i \\ 0 & \text{otherwise} \end{cases}$$

Let X be the number of people that enter on the ground floor. Then we obtain

$$E\left[\sum_{i=1}^N I_i | X = k\right] = \sum_{i=1}^N E[I_i | X = k] = N\left[1 - \left(\frac{N-1}{N}\right)^k\right]$$

where

$$\begin{aligned} E[I_i | X = k] &= P\{I_i = 1 | X = k\} \\ &= P\{\text{at least one person gets off at floor } i | X = k\} \\ &= 1 - P\{\text{no person gets off at floor } i | X = k\} \\ &= 1 - \left(\frac{N-1}{N}\right)^k \end{aligned}$$

Therefore, the expected number of stops that the elevator will make can be computed as follows.

$$\begin{aligned} E\left[\sum_{i=1}^N I_i\right] &= \sum_{k=0}^{\infty} E\left[\sum_{i=1}^N I_i | X = k\right] P\{X = k\} \\ &= N - N \sum_{k=0}^{\infty} \left(\frac{N-1}{N}\right)^k e^{-10} \cdot \frac{10^k}{k!} \\ &= N - N e^{-\frac{10}{N}} \\ &= N(1 - e^{-\frac{10}{N}}) \end{aligned}$$

7. Suppose that the expected number of accidents per week at an industrial plant is 5. Suppose also that the numbers of workers injured in each accident are independent random variables with a common mean of 2.5. If the number of workers injured in each accident is independent of the number of accidents that occur, compute the expected number of workers injured in a week.

Sol. Let N = the number of accidents in a week. $E(N) = 5$.

Let X_i = the number of workers injured in i -th accident. $E(X_i) = 2.5$.

$S = \sum_{i \leq N} X_i$ = the number of workers injured in a week.

$$\begin{aligned}
 E[S|N = n] &= E\left[\sum_{i=1}^n X_i | N = n\right] \\
 &= E\left[\sum_{i=1}^n X_i\right] \\
 &= \sum_{i=1}^n E[X_i] \\
 &= 2.5n
 \end{aligned}$$

$$ES = E[E[S|N]] = E(2.5N) = 2.5 \cdot 5 = 12.5$$

8. A certain region is inhabited by r distinct types of a certain species of insect. Each insect caught will, independently of the types of the previous catches, be of type i with probability p_i .

- (a) Compute the mean and variance of the number of insects that are caught before the first type 1 catch.

Sol. Let X denote the number of insects caught before the first type 1 catch. Then

$$P(X = x) = (1 - P_1)^x P_1$$

M1. Hence,

$$\begin{aligned}
 EX &= \sum_{x=0}^{\infty} x(1 - P_1)^x P_1 \\
 &= P_1 \sum_{x=0}^{\infty} x(1 - P_1)^x \\
 &= P_1 \cdot \frac{1 - P_1}{P_1^2} \\
 &= \frac{1 - P_1}{P_1}
 \end{aligned}$$

M2. Since $X + 1 \sim \text{Geo}(P_1)$

$$EX = \frac{1}{P_1} - 1$$

- (b) Compute the mean and variance of the number of types of insects that are caught before the first type 1 catch.

Sol. Let X_i denote the number of insects of type i caught before the first type 1 catch.

Let $g(0) = 0$ and $g(x) = 1$ for positive integers $x > 0$. We want to compute

$$E\left[\sum_{i=2}^r g(X_i)\right] = \sum_{i=2}^r E[g(X_i)] = \sum_{i=2}^r P(g(X_i) = 1) = \sum_{i=2}^r P(X_i \geq 1)$$

Let X denote the number of insects caught before the first type 1 catch as in part (a). Then

$$\begin{aligned} P(X_i \geq 1) &= \sum_{x=0}^{\infty} P(X_i \geq 1, X = x) \\ &= \sum_{x=0}^{\infty} \{P(X = x) - P(X_i = 0, X = x)\} \\ &= \sum_{x=0}^{\infty} \{(1 - P_1)^x P_1 - (1 - P_1 - P_i)^x P_1\} \\ &= \frac{P_1}{P_1} - \frac{P_1}{P_1 + P_i} \\ &= \frac{P_i}{P_1 + P_i} \end{aligned}$$

Hence,

$$E\left[\sum_{i=2}^r g(X_i)\right] = \sum_{i=2}^r P(X_i \geq 1) = \sum_{i=2}^r \frac{P_i}{P_1 + P_i}$$

9. Initially there were 10 new tennis balls and 10 used tennis balls in the bin. We choose 5 balls. Let X be the number of new balls in these chosen five balls and let p_i be the probability that we choose i new balls at this first draw, i.e., $p_i = P(X = i)$. We play a tennis game with these chosen five balls (so after playing a game all these five balls become used), and put them back into the bin. Now we choose 5 balls again from the bin and let Y be the number of new balls chosen at the second draw.

- (a) Calculate p_3 .

Sol.

$$p_3 = P(X = 3) = \frac{\binom{10}{3} \binom{10}{2}}{\binom{10+10}{5}}$$

- (b) We chose 3 new balls at the first draw, i.e., $X = 3$. What is the distribution of Y in this case?

Sol. $X = 3$ (new balls is 7, used balls is 13) Y be new balls chosen at the second draw.

$$Y \sim HG(n = 5, m = 7, N = 20)$$

- (c) We chose 3 new balls at the first draw, i.e., $X = 3$. What is the mean $E(Y|X = 3)$ and variance $Var(Y|X = 3)$ of Y in this case? Just write down your answer.

Sol.

$$E(Y|X = 3) = n \cdot \frac{m}{N} = 5 \cdot \frac{7}{20}$$

$$Var(Y|X = 3) = n \cdot \frac{m}{N} \cdot \frac{N - m}{N} \cdot \left(1 - \frac{n - 1}{N - 1}\right) = 5 \cdot \frac{7}{20} \cdot \frac{13}{20} \cdot \left(1 - \frac{4}{19}\right)$$

- (d) Represent the mean EY and variance $Var(Y)$ of Y in terms of p_i . Don't calculate p_i explicitly and don't simplify your answer. Represent your answer in terms of p_i and the summation.

Sol.

$$E(Y|X = i) = n \cdot \frac{m}{N} = 5 \cdot \frac{10 - i}{20}, \quad i = 0, \dots, 5$$

$$E(Y|X) = 5 \cdot \frac{10 - X}{20}$$

$$E(Y) = E(E(Y|X)) = \sum_{i=0}^5 5 \cdot \frac{10 - i}{20} \cdot p_i$$

$$Var(Y|X = i) = 5 \cdot \frac{10 - i}{20} \cdot \frac{10 + i}{20} \cdot \frac{15}{19}$$

$$E[Var(Y|X)] = \sum_{i=0}^5 5 \cdot \frac{10 - i}{20} \cdot \frac{10 + i}{20} \cdot \frac{15}{19} \cdot p_i$$

$$Var[E(Y|X)] = Var\left(\frac{20 - X}{4}\right) = \frac{1}{16} \cdot 5 \cdot \frac{10}{20} \cdot \frac{10}{20} \cdot \frac{15}{19}$$

Therefore,

$$\begin{aligned} VarY &= E[Var(Y|X)] + Var[E(Y|X)] \\ &= \sum_{i=0}^5 5 \cdot \frac{10 - i}{20} \cdot \frac{10 + i}{20} \cdot \frac{15}{19} \cdot p_i + \frac{1}{16} \cdot 5 \cdot \frac{10}{20} \cdot \frac{10}{20} \cdot \frac{15}{19} \end{aligned}$$

10. Let's flip the fair coin until we get the pattern HH . Let W_{HH} be the waiting time to get the pattern HH .

(a) Calculate EW_{HH} .

Sol. Let X denote the number of flips the fair coin until we have the first head. Then $X \sim \text{Geo}(\frac{1}{2})$.

Now we renew to flip the fair coin until we see the first tail and call the number of such extra flips Y . Then $Y \sim \text{Geo}(\frac{1}{2})$.

Since X and Y are independent,

$$EW_{HT} = EX + EY = \frac{1}{\frac{1}{2}} + \frac{1}{\frac{1}{2}} = 4$$

$$E(W_{HH}|H) = P(H)E(W_{HH}|HH) + P(T)E(W_{HH}|HT) = P(H) \cdot 2 + P(T) \cdot (2 + EW_{HH})$$

$$E(W_{HH}|T) = 1 + EW_{HH}$$

$$\begin{aligned} EW_{HH} &= P(H)E(W_{HH}|H) + P(T)E(W_{HH}|T) \\ &= \frac{1}{2}\{1 + \frac{1}{2}(2 + EW_{HH})\} + \frac{1}{2}(1 + EW_{HH}) \end{aligned}$$

$$EW_{HH} = 6$$

(b) Calculate $\text{Var}(W_{HH})$.

Sol.

$$\begin{aligned} E(W_{HH}^2|H) &= P(H)E(W_{HH}^2|HH) + P(T)E(W_{HH}^2|HT) \\ &= \frac{1}{2} \cdot 2^2 + \frac{1}{2}E(W_{HH} + 2)^2 \\ &= \frac{1}{2}EW_{HH}^2 + 16 \end{aligned}$$

$$\begin{aligned} E(W_{HH}^2|T) &= E((W_{HH} + 1)^2) \\ &= EW_{HH}^2 + 2EW_{HH} + 1 \\ &= EW_{HH}^2 + 13 \end{aligned}$$

$$EW_{HH}^2 = P(H)E(W_{HH}^2|H) + P(T)E(W_{HH}^2|T) = 58$$

Therefore

$$\text{Var}W_{HH} = EW_{HH}^2 - (EW_{HH})^2 = 22$$

[Extra]

1. Let X be the number of 1's and Y the number of 2's that occur in n rolls of a fair die. Compute $Cov(X, Y)$.
2. Roll the fair dice twice independently. X is the sum of the two and Y is the number of 2 appears. Calculate $Cov(X, Y)$.
3. A coin having probability p of coming up heads is continually flipped until both heads and tails have appeared. Find the expectation and variance of the number of flips.
4. A fair die is successively rolled. Let X be the number of rolls to get the first 6 and let Y be the number of rolls to get the first 5. Find
 - (a) $E[X]$.
 - (b) $E[X|Y = 1]$.
 - (c) $E[X|Y = 5]$.
 - (d) $Var[X|Y = 5]$.
5. Consider a population consisting of individuals able to produce offspring of the same kind. Suppose that, by the end of its lifetime, each individual will have produced j new offspring with probability p_j , $j \geq 0$, independently of the number produced by any other individual. The number of individuals initially present, denoted by X_0 , is called the size of the zero-th generation. All offspring of the zero-th generation constitute the first generation, and their number is denoted by X_1 . In general, let X_n denote the size of the n -th generation. Let μ and σ^2 be the mean and the variance of the number of offspring produced by a single individual;

$$\mu = \sum_{j=0}^{\infty} jp_j \quad \text{and} \quad \sigma^2 = \sum_{j=0}^{\infty} (j - \mu)^2 p_j.$$

Suppose that $X_0 = 1$, that is, initially there is a single individual in the population.

- (a) Show that $EX_n = \mu EX_{n-1}$.
- (b) Use (a) to conclude $EX_n = \mu^n$.
- (c) Show that $Var(X_n) = \sigma^2 \mu^{n-1} + \mu^2 Var(X_{n-1})$.

(d) Use (c) to conclude

$$\text{Var}(X_n) = \begin{cases} \sigma^2 \mu^{n-1} \left(\frac{\mu^n - 1}{\mu - 1} \right) & \text{if } \mu \neq 1, \\ n\sigma^2 & \text{if } \mu = 1. \end{cases}$$