Stat 88: Probability & Mathematical Statistics in Data Science

Lecture 11: 2/22/2022

Sums of Poisson rvs, Expectation, Functions of rvs, Indicators

Sections 5.1, 5.2, 5.3 & 4.4

Shobhana M Stoyanov

Recall the Poisson Distribution

- Used to model rare events. X is the number of times a rare event occurs, $X = 0, 1, 2, \ldots$

- We say that a random variable X has the ***Poisson*** distribution if

$$\boxed{\boldsymbol{P(X = k) = e^{-\mu}\frac{\mu^k}{k!}}}, \quad X = 0, 1, 2, \ldots$$

- The parameter of the distribution is μ

$X \sim \text{Pois}(\mu)$

$$P(X=0) = e^{-\mu}$$

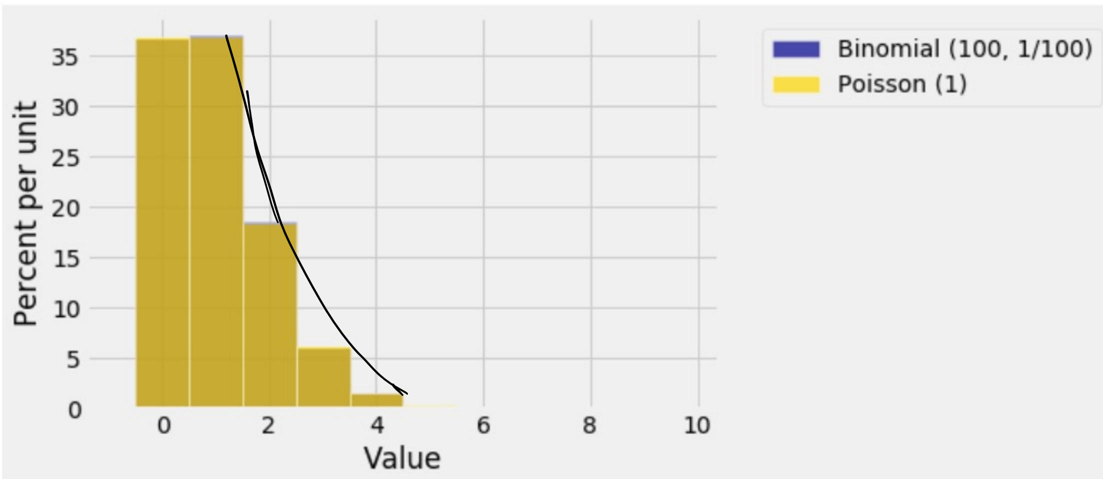
$$P(X=1) = \mu e^{-\mu}$$

$$P(X=2) = e^{-\mu}\frac{\mu^2}{2}$$

$$P(X=3) = e^{-\mu}\frac{\mu^3}{6}$$

Relationship between Poisson and Binomial distributions

- ***The Law of Small Numbers***: when $\boldsymbol{n}$ is large and $\boldsymbol{p}$ is small, the binomial (n,p) distribution is ***well approximated*** by the Poisson(μ) distribution where $\boldsymbol{\mu = np}$.



Exercise 4.5.7

$X_k =$ # of misprints in a chapter k
$X_k \sim Pois(2)$ for $1 \le k \le 20$

A book has 20 chapters. In each chapter the number of misprints has the Poisson distribution with parameter 2, independently of the misprints in other chapters.

a) Find the chance that Chapter 1 has more than two misprints.

b) Find the chance that the book has no misprints.

c) Find the chance that two of the chapters have three misprints each.

$$(a)\ P(X_1 > 2) = 1 - P(X_1 \le 2) = 1 - [P(X_1 = 0) + P(X_1 = 1) + P(X_1 = 2)]$$
$$= 1 - \left[e^{-2} + 2e^{-2} + e^{-2} \cdot \frac{2^2}{2!}\right]$$

$$(b)\ P(\text{no misprints in any chapter}) = P(X_1 = 0 \ \&\ X_2 = 0 \ \& \cdots X_{20} = 0)$$
$$= (e^{-2})^{20}$$

$$(c)\ P(X_k = 3) = e^{-2} \cdot \frac{2^3}{3!} = p = P(3 \text{ misprints in a chapter})$$

Y = # of chapters w/ 3 misprints

$$Y \sim Bin(20, p), \quad P(Y = 2) = \binom{20}{2} p^2 (1-p)^{18}$$

Sums of independent Poisson random variables

- If X and Y are random variables such that
- X and Y are independent,
- X has the Poisson(μ) distribution, and
- Y has the Poisson(λ) distribution,
- then the sum $S=X+Y$ has the Poisson ($\mu+\lambda$) distribution.

$\gamma = \mu + \lambda$

$$P(X+Y=k) = \frac{e^{-\gamma} \cdot \gamma^{k}}{k!}$$

Exercise 4.5.8

In the first hour that a bank opens, the customers who enter are of three kinds: those who only require teller service, those who only want to use the ATM, and those who only require special services (neither the tellers nor the ATM). Assume that the numbers of customers of the three kinds are independent of each other, and also that:

- the number that only require teller service has the Poisson (6) distribution, $X \sim Pois(6)$
- the number that only want to use the ATM has the Poisson (2) distribution, and $Y \sim Pois(2)$
- the number that only require special services has the Poisson (1) distribution. $Z \sim Pois(1)$

Suppose you observe the bank in the first hour that it opens. In each part below, find the chance of the event described.

$W = X + Y + Z \sim Pois(9)$

a) 12 customers enter the bank

b) more than 12 customers enter the bank

c) customers do enter but none requires special services

(a) $P(W = 12) = e^{-9}\frac{9^{12}}{12!}$

cdf

(b) $P(W > 12) = 1 - P(W \le 12) = 1 - F(12)$

$= 1 - \sum_{k=0}^{12} e^{-9}\frac{9^k}{k!}$

(c) $P(Z=0 \;\&\; X+Y>0) = P(Z=0)\,P(X+Y>0)$

$$= e^{-1}(1 - P(X+Y=0))$$

$$= e^{-1}(1 - e^{-8})$$

5.1: Expected Value of a random variable

- The ***Expectation*** or ***Expected Value*** of a random variable X is defined to be the sum of all the products $(x \times f(x))$ over all possible values x of the random variable X: that is, the expectation of a random variable is a ***weighted average*** of all the possible values that the rv can take, weighted by the probability of each value.

- $E(X) = \sum_x x \cdot P(X = x) = \sum_x x \cdot f(x)$

$P(H) = \frac{1}{2}$

- For example, toss a fair coin 3 times, let X = # of heads. Write down $f(x)$, and then use the formula to compute the expectation of X.

x	0	1	2	3
$f(x)$	1/8	3/8	3/8	1/8

$$E(X) = 0 \cdot \frac{1}{8} + 1 \cdot \frac{3}{8} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8}$$

$$= \frac{0 + 3 + 6 + 3}{8} = \frac{12}{8} = 1.5$$

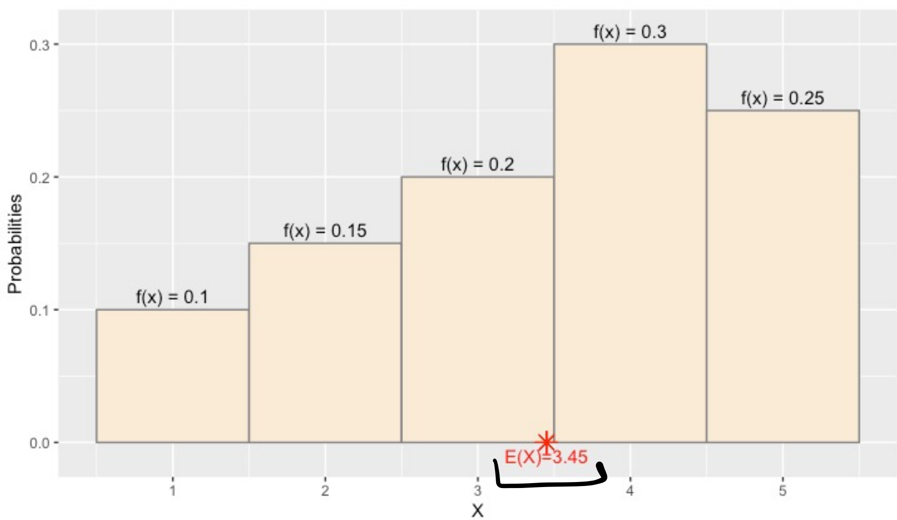
- Note: If X takes finitely many values, no problem. If it takes infinitely (countable, but not finite) many values (such as Poisson, or Geometric), then we have to be more careful when we take the sums.

Example (from text)

$$\mathbb{E}(X) = \sum_x x \cdot P(X=x)$$

x	1	2	3	4	5
f(x)	0.1	0.15	0.2	0.3	0.25

$$= 1(0.1) + 2(0.15) + 3(0.2) + 4(0.3) + 5(0.25)$$



$\mathbb{E}(X)$ as the center of mass of the prob. dsn.

Notes about E(X)

- Same units as X
- Not necessarily an attainable value (for example, in In 2020, there was an average of 1.93 children under 18 per family in the United States)
- Expectation is a long-run average value of X
- Center of mass or center of gravity (balancing point) for the distribution
- Expectation of a constant is that constant $E(c) = c$
- Linearity: $\boxed{E(aX + b) \ = aE(X) \ + b}$
- Example: for the X defined in the previous slide, find E(2X-1)

$\sum_x (ax+b) f(x)$

$Y = aX + b$

$E(Y)$

$\mathbb{E}(aX+b) = \sum_x (ax+b) \overbrace{P(X=x)}^{f(x)}$

$= \sum_x ax f(x)$

$+ \sum_x b \cdot f(x)$

$= a \sum_x x f(x) + b \sum f(x)$

$a \cdot \mathbb{E}(X) + b \cdot 1$

$\mathbb{E}(X) = 3.45$

$\mathbb{E}(2X-1) = 2\mathbb{E}(X) - 1$

$= 2 \times 3.45 - 1$
$= 5.9$

Special examples

- Bernoulli (Indicators)

$X \sim \text{Bernoulli}(p)$ if $X = \begin{cases} 1 & \text{with prob } p \\ 0 & \text{w.p. } 1-p \end{cases}$

$$\mathbb{E}(X) = 1 \cdot p + 0 \cdot (1-p) = p$$

Let A be an event with prob. p. $[P(A) = p]$

Define r.v. $I_A = \begin{cases} 1 & \text{if } A \text{ is true} \\ 0 & \text{o/w} \end{cases}$

$$\mathbb{E}(I_A) = 1 \cdot P(I_A = 1) + 0 \cdot P(I_A = 0) = P(I_A = 1) = P(A)$$

- discrete Uniform $X \sim \text{unif}\{1, 2, \dots, n\}$

$P(X = x) = \frac{1}{n}$

$$\mathbb{E}(X) = 1 \cdot \frac{1}{n} + 2 \cdot \frac{1}{n} + 3 \cdot \frac{1}{n} + \dots + n \cdot \frac{1}{n} = \frac{1+2+3+\dots+n}{n}$$

$$= \frac{1}{n}[1+2+\dots+n]$$

$$= \frac{1}{\cancel{n}}\left[\frac{\cancel{n}(n+1)}{2}\right] = \frac{n+1}{2}$$

- Poisson

$$P(X = k) = e^{-\mu} \frac{\mu^k}{k!}, \quad k = 0, 1, 2, \dots$$

$$\mathbb{E}(X) = \sum_x x \cdot P(X=x) = \sum_{k=0}^{\infty} k \cdot e^{-\mu} \cdot \frac{\mu^k}{k!}$$

$$= \sum_{k=1}^{\infty} k \cdot e^{-\mu} \cdot \frac{\mu^k}{k!} = e^{-\mu} \sum_{k=1}^{\infty} \cancel{k} \cdot \frac{\mu^k}{\cancel{k!}\,(k-1)!}$$

$$\mathbb{E}(X) = e^{-\mu} \sum_{k=1}^{\infty} \frac{\mu^k}{(k-1)!} = e^{-\mu} \cdot \mu \sum_{k=1}^{\infty} \frac{\mu^{k-1}}{(k-1)!}$$

Let $j = k-1$, so j goes from 0 to ∞

$$\mathbb{E}(X) = e^{-\mu} \cdot \mu \cdot \underbrace{\sum_{j=0}^{\infty} \frac{\mu^j}{j!}}_{= e^{\mu}}$$

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \dots$$

$$\mathbb{E}(X) = e^{-\mu} \cdot \mu \cdot e^{+\mu} = \mu$$

If $X \sim \text{Pois}(\mu)$, $\mathbb{E}(X) = \mu$

5.2: Functions of random variables

$\mathbb{E}(Y) = \sum_y y \cdot P(Y=y)$

- X~unif{-1,0,1}, $Y = X^2$, find $E(Y)$

x	-1	0	1
$f(x)$	$1/3$	$1/3$	$1/3$
y	$(-1)^2=1$	$0^2=0$	$1^2=1$

$= 1 \cdot \frac{2}{3} + 0 \cdot \frac{1}{3} = \frac{2}{3}$

$\mathbb{E}(Y) = (-1)^2 \cdot \frac{1}{3} + 0^2 \cdot \frac{1}{3} + 1^2 \cdot \frac{1}{3} = 2/3$

- In general, if $\boldsymbol{Y = g(X)}$, $\boldsymbol{E(Y) = E\big(g(X)\big) = \sum g(x) \cdot f(x)} = \sum \boldsymbol{g(x) \cdot P(X = x)}$

$\mathbb{E}(Y) = \sum_x x^2 \cdot f(x)$

- Example: $W = \min(X, 0.5)$. $Find\ E(W)$ (write out the values of W, and probs)

x	-1	0	1
$f(x)$	$1/3$	$1/3$	$1/3$
w	-1	0	$1/2$

$\mathbb{E}(W) = (-1)\frac{1}{3} + 0 \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{3}$

$= -\frac{1}{2} \cdot \frac{1}{3}$

- Note that $Y = g(X)$ is also a random variable, just defined via X.

Say $Y = aX + b$ $\qquad$ $\mathbb{E}(Y) = \mathbb{E}(aX+b)$

$Y = g(X)$ $\qquad$ $= \sum_x (ax+b) \cdot f(x)$

$\mathbb{E}(Y) = \sum_x g(x) f(x)$

$X+Y$

Multiple random variables on the same outcome space

- **Joint distributions**: Recall rolling a *pair* of dice. Draw a table of outcomes and their probabilities. What event does each cell represent?
- Sum of probabilities across *all* the cells is 1, since this is all the probabilities across all possible outcomes

X_1 : outcome of 1^{st} die
X_2 " " 2^{nd} die

$X_1 \backslash X_2$	1	2	3	4	5	6
1	(1,1)	(1,2)	(1,3)	. . .		
2						
3						
4						
5						
6						

each outcome is equally likely w.p $\frac{1}{36}$

Joint distributions

1 2 2 3 3

- Draw two tickets without replacement from a box that has 5 tickets marked 1, 2, 2, 3, 3. Let X_1 and X_2 represent the values of the tickets drawn on the first and second draws respectively.
- Create a table of all possible outcomes for the pair (X_1, X_2) (which is also a random variable), and write down the probabilities using the multiplication rule.

Write probs of (X_1, X_2) in cell $P(X_1=k \,\&\, X_2=j)$

	$X_2 = 1$	$X_2 = 2$	$X_2 = 3$	marginal dsn of X_1 ↓
$X_1 = 1$	$P(X_1=1 \,\&\, X_2=1) = 0$	$P(X_1=1 \,\&\, X_2=2) = P(X_2=2 \mid X_1=1)P(X_1=1) = (1/5)(1/2) = 1/10$	$P(X_2=3 \,\&\, X_1=1) = P(X_1=1)P(X_2=3 \mid X_1=1) = 1/10$	$P(X_1=1) = \frac{1}{5}$
$X_1 = 2$	$1/10 = \frac{2}{20}$	$P(X_1=2 \,\&\, X_2=2) = P(X_1=2)P(X_2=2 \mid X_1=2)$ $(2/5)(1/4) = 2/20$	$4/20$	$P(X_1=2) = 8/20 = \frac{2}{5}$
$X_1 = 3$	$1/10 = \frac{2}{20}$	$4/20$	$2/20$	$\frac{8}{20} = \frac{2}{5} = P(X_1=3)$
Marginal distributions of X_2 →	$P(X_1=1) = \frac{4}{20}$	$P(X_2=2) = \frac{8}{20}$	$P(X_2=3) = \frac{8}{20}$	

Marginal distributions

- What is $\mathrm{P}(X_1 = 1)$? Write down the pmf for X_1 and X_2

- Are they independent?

- Use the table to compute $P(X_1 + X_2 = 5)$

- Use the table to compute $E\big(g(X_1, X_2)\big)$, where $g(X_1, X_2) = | X_1 - X_2 |$ (the expected distance between the two draws)

- If $S = X_1 + X$_2, find $E(S)$.

Method of indicators

- Additivity of Expectation: This is a very useful property – no matter what the joint distribution of X and Y may be, we have:
$$E(X+Y) = E(X) + E(Y)$$
- Whether X and Y are dependent or independent, this holds, making it enormously useful.
- We also have linearity: $E(aX + bY) = aE(X) + bE(Y)$

- Recall that we talked about "classifying and counting" – that is, we divide up the outcomes into those that we are interested in (successes), and everything else (failures), and then count the number of successes.
- We can represent these outcomes as 0 and 1, where 1 marks a success and 0 and failure, so if we model the trials as draws from a box, we can count the number of success by counting up the number of times we drew a 1.

- We can represent each draw as a Bernoulli trial, where $p = P(S)$

Using indicators and additivity

- For example, say we roll a die 10 times, and success is rolling a 1.
- Then $p = 1/6$, and we can define a Bernoulli rv as $X = \begin{cases} 0, & w.p. 5/6 \\ 1, & w.p. 1/6 \end{cases}$
- We can also define an event A: let A be the event of rolling a 1 and define a rv I_A that takes the value 1 if A occurs and 0 otherwise.
- This is a Bernoulli rv, what is its expectation?
- Now let $X \sim Bin(10, \frac{1}{6})$, so X counts the number of successes in 10 rolls. Let's find $E(X)$ using additivity and indicators:

Using indicators

- Binomial

- Hypergeometric: Did we use the independence of the trials for the binomial? If not, we can use the same method to compute the expected value of a hypergeometric rv:

Exercise 5.7.6: A die is rolled 12 times. Find the expectation of:

a) the number of times the face with five spots appears

b) the number of times an odd number of spots appears

c) the number of faces that don't appear

d) the number of faces that do appear

Example

- Let X be the number of spades in 7 cards dealt **with replacement** from a well shuffled deck of 52 cards containing 13 spades. Find E(X).

1. Write down what X is

2. Define an indicator for the kth trial: I_k

3. Find $p = P(I_k = 1)$

4. Write X as a sum of indicators

5. Now compute $E(X)$ using additivity

- Do the same thing if we deal 7 cards **without replacement.**

Missing classes

- We can use indicators to compute the chance that something *doesn't* occur.

- For example, say we have a box with balls that are red, white, or blue, with 35% being red, 30% being white, and 35% blue. If we draw n times with replacement from this box, what is the expected number of colors that *don't* appear in the sample?

Examples

1. An instructor is trying to set up office hours during RRR week. On one day there are 8 available slots: 10-11, 11-noon, noon-1, 1-2, 2-3, 3-4, 4-5, and 5-6. There are 6 GSIs, each of whom picks one slot. Suppose the GSIs pick the slots at random, independently of each other. Find the expected number of slots that no GSI picks.

2. A building has 10 floors above the basement. If 12 people get into an elevator at the basement, and each chooses a floor at random to get out, independently of the others, at how many floors do you expect the elevator to make a stop to let out one or more of these 12 people?