

Wrap Poisson Distribution + Problems

CSE 312 Spring 26
Lecture 14

Announcements

- No concept check today
- Material for midterm ends today.
- Midterm covers
 - everything from class up till and including today
 - All concept checks through today
 - Problem Sets 1-5
 - Sections 1-5
- I will post practice tests tomorrow.

Discrete Uniform Random Variables

A discrete random variable X **equally likely** to take any (integer) value between integers a and b (inclusive), is **uniform**.

Notation: $X \sim \text{Unif}(a, b)$

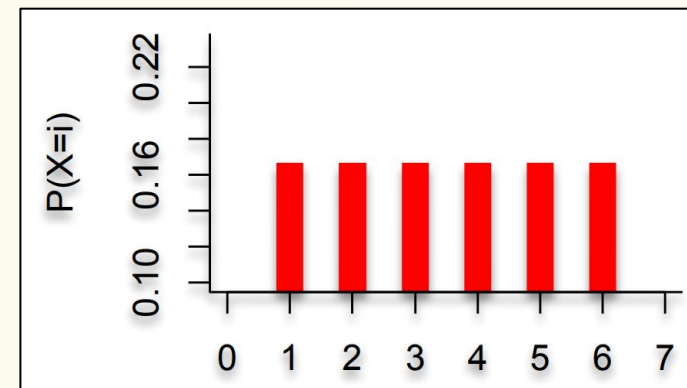
PMF: $P(X = i) = \frac{1}{b - a + 1}$

Expectation: $\mathbb{E}[X] = \frac{a+b}{2}$

Variance: $\text{Var}(X) = \frac{(b-a)(b-a+2)}{12}$

Example: value shown on one roll of a fair die is $\text{Unif}(1,6)$:

- $P(X = i) = 1/6$
- $\mathbb{E}[X] = 7/2$
- $\text{Var}(X) = 35/12$



Bernoulli Random Variables

A random variable X that takes value **1** (“Success”) with probability p , and **0** (“Failure”) otherwise. X is called a **Bernoulli random variable**.

Notation: $X \sim \text{Ber}(p)$

PMF: $P(X = 1) = p, P(X = 0) = 1 - p$

Expectation: $\mathbb{E}[X] = p$ Note: $\mathbb{E}[X^2] = p$

Variance: $\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = p - p^2 = p(1 - p)$

Examples:

- Coin flip
- Randomly guessing on a MC test question
- A server in a cluster fails
- Whether or not a share of a particular stock pays off or not
- Any indicator r.v.

Binomial Random Variables

A discrete random variable $X = \sum_{i=1}^n Y_i$ where each $Y_i \sim \text{Ber}(p)$.

Counts number of successes in n independent trials, each with probability p of success.

X is a **Binomial random variable**

Notation: $X \sim \text{Bin}(n, p)$

PMF: $P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$

Expectation: $\mathbb{E}[X] = np$

Variance: $\text{Var}(X) = np(1 - p)$

Examples:

- # of heads in n indep coin flips
- # of 1s in a randomly generated n bit string
- # of servers that fail in a cluster of n computers
- # of bit errors in file written to disk
- # of elements in a bucket of a large hash table
- # of n different stocks that “pay off”

Geometric Random Variables

A discrete random variable X that models the number of independent trials $Y_i \sim \text{Ber}(p)$ up until and including the first success.

X is called a **Geometric random variable** with parameter p .

Notation: $X \sim \text{Geo}(p)$

PMF: $P(X = k) = (1 - p)^{k-1}p$

Expectation: $\mathbb{E}[X] = \frac{1}{p}$

Variance: $\text{Var}(X) = \frac{1-p}{p^2}$

Examples:

- # of coin flips until first head
- # of random guesses on MC questions until you get one right
- # of random guesses at a password until you hit it
- # hash trials until a miner successfully mines a Bitcoin

Poisson Distribution

- Suppose “events” happen, independently, at an *average* rate of λ per unit time.
- Let X be the *actual* number of events happening in a given time unit. Then X is a *Poisson* r.v. with parameter λ (denoted $X \sim \text{Poi}(\lambda)$) and has distribution (PMF):

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

Some examples of “Poisson processes”:

- # of cars passing through a traffic light in 1 hour
 - # of requests to web servers in an hour
 - # of photons hitting a light detector in a given interval
 - # of patients arriving to ER within an hour
- Assume fixed average rate

Poisson Distribution – summary (1)

Siméon Denis Poisson
1781-1840

- Suppose “events” happen, independently, at an *average* rate of λ per unit time.
- Let X be the *actual* number of events happening in a given time unit. Then X is a *Poisson* r.v. with parameter λ (denoted $X \sim \text{Poi}(\lambda)$) and has distribution (PMF):

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

$$\begin{aligned} E(X) &= \lambda \\ \text{Var}(X) &= \lambda \end{aligned}$$


From Binomial to Poisson

$$X \sim \text{Bin}(n, p)$$

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

$$E[X] = np$$

$$\text{Var}(X) = np(1 - p)$$

$$\begin{aligned} n &\rightarrow \infty \\ np &= \lambda \\ p &= \frac{\lambda}{n} \rightarrow 0 \end{aligned}$$


$$X \sim \text{Poi}(\lambda)$$

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

$$E[X] = \lambda$$

$$\text{Var}(X) = \lambda$$

Sum of Independent Poisson RVs

Let $X \sim \text{Poi}(\lambda_1)$ and $Y \sim \text{Poi}(\lambda_2)$ such that $\lambda = \lambda_1 + \lambda_2$.

Let $Z = X + Y$. What kind of random variable is Z ?

Aka what is the “distribution” of Z ?

Intuition first:

- X is measuring number of (type 1) events that happen in, say, an hour if they happen at an average rate of λ_1 per hour.
- Y is measuring number of (type 2) events that happen in, say, an hour if they happen at an average rate of λ_2 per hour.
- Z is measuring total number of events of both types that happen in, say, an hour, if type 1 and type 2 events occur independently.

Sum of Independent Poisson RVs (2)

Theorem. Let $X \sim \text{Poi}(\lambda_1)$ and $Y \sim \text{Poi}(\lambda_2)$ be independent r.v.s such that $\lambda = \lambda_1 + \lambda_2$. Let $Z = X + Y$. For all $z = 0, 1, 2, 3 \dots$,

$$P(Z = z) = e^{-\lambda} \cdot \frac{\lambda^z}{z!}$$

More generally, let $X_1 \sim \text{Poi}(\lambda_1), \dots, X_n \sim \text{Poi}(\lambda_n)$ independent such that $\lambda = \sum_i \lambda_i$. Let $Z = \sum_i X_i$

$$P(Z = z) = e^{-\lambda} \cdot \frac{\lambda^z}{z!}$$

Sum of Independent Poisson RVs (3)

Theorem. Let $X \sim \text{Poi}(\lambda_1)$ and $Y \sim \text{Poi}(\lambda_2)$ such that $\lambda = \lambda_1 + \lambda_2$.

Let $Z = X + Y$. For all $z = 0, 1, 2, 3, \dots$,

$$P(Z = z) = e^{-\lambda} \cdot \frac{\lambda^z}{z!}$$

$P(Z = z) = ?$

1. $P(Z = z) = \sum_{j=0}^z P(X = j, Y = z - j)$
2. $P(Z = z) = \sum_{j=0}^{\infty} P(X = j, Y = z - j)$
3. $P(Z = z) = \sum_{j=0}^z P(Y = z - j | X = j) P(X = j)$
4. $P(Z = z) = \sum_{j=0}^z P(Y = z - j | X = j)$

<https://pollev.com/annakarlin185>

- A. All of them are right
- B. The first 3 are right
- C. Only 1 is right
- D. Don't know

Law of Total Probability

Law of Total Probability – the details

Theorem: [Law of Total Probability for Discrete R.V.s]

Let E be an event. Let Y be a discrete r.v.

$$P\{E\} = \sum_y P\{E \cap Y = y\} = \sum_y P\{E | Y = y\} \cdot P\{Y = y\}$$

For a discrete r.v. X :

$$P\{X = k\} = \sum_y P\{X = k \cap Y = y\} = \sum_y P\{X = k | Y = y\} \cdot P\{Y = y\}$$

Proof: Follows immediately from Law of Total Probability for Events, if we realize that $Y = y$ represents an event and the set of events $Y = y$ over all y form a partition.

Theorem. Let $X \sim \text{Poi}(\lambda_1)$ and $Y \sim \text{Poi}(\lambda_2)$ such that $\lambda = \lambda_1 + \lambda_2$.

Let $Z = X + Y$. For all $z = 0, 1, 2, 3 \dots$,

$$P(Z = z) = e^{-\lambda} \cdot \frac{\lambda^z}{z!}$$

Proof

$$P(Z = z) = \sum_{j=0}^z P(X = j, Y = z - j) \quad \text{Law of total probability}$$

Proof. Written out

$$P(Z = z) = \sum_{j=0}^z P(X = j, Y = z - j)$$

Law of total probability

$$= \sum_{j=0}^z P(X = j) P(Y = z - j) = \sum_{j=0}^z e^{-\lambda_1} \cdot \frac{\lambda_1^j}{j!} \cdot e^{-\lambda_2} \cdot \frac{\lambda_2^{z-j}}{(z-j)!} \quad \text{Independence}$$

$$= e^{-\lambda_1 - \lambda_2} \left(\sum_{j=0}^z \frac{1}{j! (z-j)!} \cdot \lambda_1^j \lambda_2^{z-j} \right)$$

$$= e^{-\lambda} \left(\sum_{j=0}^z \frac{z!}{j! (z-j)!} \cdot \lambda_1^j \lambda_2^{z-j} \right) \frac{1}{z!}$$

$$= e^{-\lambda} \cdot (\lambda_1 + \lambda_2)^z \cdot \frac{1}{z!} = e^{-\lambda} \cdot \lambda^z \cdot \frac{1}{z!}$$

Binomial
Theorem

Poisson Random Variables - Summary

Definition. A **Poisson random variable** X with parameter $\lambda \geq 0$ is such that for all $i = 0, 1, 2, 3 \dots$,

$$P(X = i) = e^{-\lambda} \cdot \frac{\lambda^i}{i!}$$

General principle:

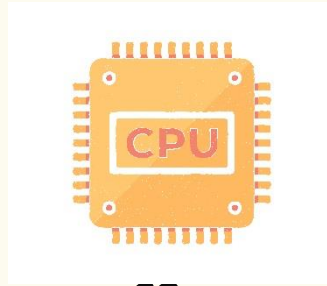
- Events happen at an average rate of λ per time unit
- Number of events happening at a time unit X is distributed according to $\text{Poi}(\lambda)$
- Poisson approximates Binomial when n is large, p is small, and np is moderate
- Sum of independent Poisson is still a Poisson

Who Fails First?

You have a disk with probability p_1 of failing each day, and a CPU which independently has probability p_2 of failing each day.



p_1



p_2

Q: What is the probability that the disk fails *before* the CPU?

Who Fails First? (1)

Disk with prob. p_1 of failing each day, and a CPU with indpt. prob. p_2 of failing each day.

Q: What is the probability that the disk fails *before* the CPU? (Redo using conditioning!)

$X_1 =$ days until disk fails $\sim \text{Geometric}(p_1)$

$X_2 =$ days until CPU fails $\sim \text{Geometric}(p_2)$

$X_1 \perp X_2$

$\mathbf{P}\{X_1 < X_2\} =$

Who Fails First? (2)

Disk with prob. p_1 of failing each day, and a CPU with indpt. prob. p_2 of failing each day.

Q: What is the probability that the disk fails *before* the CPU? (Redo using conditioning!)

$X_1 =$ days until disk fails $\sim \text{Geometric}(p_1)$

$X_2 =$ days until CPU fails $\sim \text{Geometric}(p_2)$

$$\mathbf{P}\{X_1 < X_2\} = \sum_{k=1}^{\infty} \mathbf{P}\{X_1 < X_2 \mid X_1 = k\} \cdot \mathbf{P}\{X_1 = k\}$$

$X_1 \perp X_2$

$$= \sum_{k=1}^{\infty} \mathbf{P}\{k < X_2 \mid X_1 = k\} \cdot \mathbf{P}\{X_1 = k\}$$

$$= \sum_{k=1}^{\infty} \mathbf{P}\{X_2 > k\} \cdot \mathbf{P}\{X_1 = k\}$$

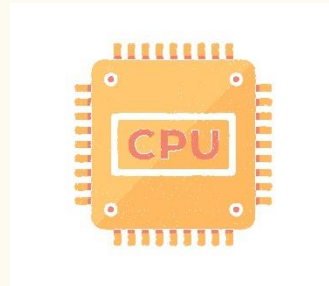
$$= \sum_{k=1}^{\infty} (1 - p_2)^k \cdot (1 - p_1)^{k-1} \cdot p_1 = \frac{p_1(1 - p_2)}{1 - (1 - p_2)(1 - p_1)}$$

Who Fails First? (3)

You have a disk with probability p_1 of failing each day, and a CPU which independently has probability p_2 of failing each day.



p_1



p_2

$$P\{\text{disk fails before CPU fails}\} = \frac{p_1(1-p_2)}{1-(1-p_2)(1-p_1)}$$

But WHY?

Who Fails First? (4)

You have a disk with probability p_1 of failing each day, and a CPU which independently has probability p_2 of failing each day.

$$P\{\text{disk fails before CPU fails}\} = \frac{p_1(1-p_2)}{1-(1-p_2)(1-p_1)}$$

Intuition: Think about flipping 2 coins each day.

There may be many days where both coins are heads.

We only care about the *first day where the coins are not both heads*.

Given that both coins are not heads, what's the probability that coin 1 is H and coin 2 is T?

$$P\{\text{coin 1 is H \& coin 2 is T} \mid \text{not both tails}\} =$$

Who Fails First? (5)

You have a disk with probability p_1 of failing each day, and a CPU which independently has probability p_2 of failing each day.

$$P\{\text{disk fails before CPU fails}\} = \frac{p_1(1-p_2)}{1-(1-p_2)(1-p_1)}$$

Intuition: Think about flipping 2 coins each day.

There may be many days where both coins are heads.

We only care about the *first day where the coins are not both heads*.

Given that both coins are not heads, what's the probability that coin 1 is H and coin 2 is T?

$$P\{\text{coin 1 is H \& coin 2 is T} \mid \text{not both tails}\} = \frac{P\{\text{coin 1 is H \& coin 2 is T}\}}{P\{\text{not both tails}\}} = \frac{p_1(1-p_2)}{1-(1-p_2)(1-p_1)}$$

Parallel Server Failures (on section 5 worksheet)

A computing system relies on m independent components. The lifetime (in days) of each component is modeled by a Geometric distribution with parameter p (has probability p of failing each day). The system completely shuts down only when all m nodes have failed.

Let D be the day the first component fails. Find the probability mass function of D . Start by computing the probability that $D > d$.

Parallel Server Failures (on section 5 worksheet)

A computing system relies on m independent components. The lifetime (in days) of each component is modeled by a Geometric distribution with parameter p (has probability p of failing each day). The system completely shuts down only when all m nodes have failed.

Find the probability that the entire system shuts down on day k .

Homework grading (on section 5 worksheet)

Homeworks are graded by TA1 with probability 0.3, by TA2 with probability 0.5 and by TA3 with probability 0.2

TA1 takes an amount of time $X \sim \text{Poi}(\lambda)$, to grade.

TA2 takes an amount of time $Y \sim \text{Bin}(n, p)$ to grade.

TA3 takes an amount of time $Z \sim \text{Geom}(p)$ to grade.

What is the probability that TA1 did the grading given that the number of hours it took was h ?

End of material for midterm!!!

Next time: continuous random variables!