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

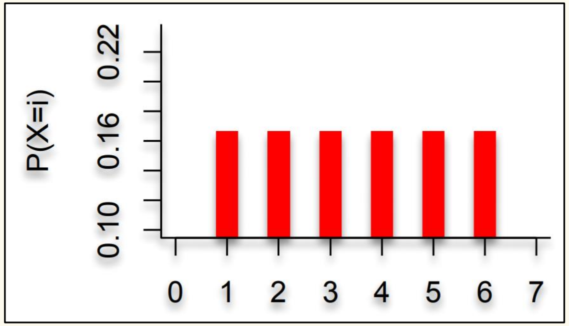
PMF: $\mathrm{P}(X = i) = \frac{1}{b - a + 1}$

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

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

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

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



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 $Y_i \sim \mathrm{Ber}(p)$.

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

X is a Binomial random variable

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"

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

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

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

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

$F_X(k) = P(X \le k) = \sum_{i=0}^{k} \binom{n}{i} p^i (1-p)^{n-i}$

Geometric Random Variables

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

X is called a Geometric random variable with parameter p.

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

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

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

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

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

$$F_X(k) = \sum_{i=1}^{k} (1-p)^{i-1} p$$

$$=$$

$$P(X \le k)$$

$$=$$

$$P(X \le k) = 1 - (1-p)^k$$

$$P(X > k) = (1-p)^k$$

$$P(X > k) = \alpha^k$$

Poisson Distribution

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 $X \sim \mathrm{Poi}(\lambda)$) and has distribution (PMF):

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

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 $X \sim \mathrm{Poi}(\lambda)$) and has distribution (PMF):

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

$$E(X) = \lambda$$
$$\mathrm{Var}(X) = \lambda$$

From Binomial to Poisson

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

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

$$E[X] = np$$

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

$$n \to \infty$$

$$np = \lambda$$

$$p = \frac{\lambda}{n} \to 0$$

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

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

$$E[X] = \lambda$$

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

Sum of Independent Poisson RVs

X, Y indep.

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 $z = 0,1,2,3 \ldots,$

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

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

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

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 $z = 0,1,2,3 \ldots,$

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

$P(X+Y=z)$

$P(Z = z) = ?$

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

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

$F_1, F_2, \ldots F_n$ partition sample space

$$P(E) = \sum_{i=1}^{n} P(E \cap F_i) = \sum_{i=1}^{n} P(E \mid F_i) P(F_i)$$

Y r.v. $\Omega_Y = \{y_1, y_2, \ldots, y_k\}$

$\{Y = y_1\}, \{Y = y_2\}, \ldots \{Y = y_k\}$ partition

$$P(E) = \sum_{i=1}^{k} P(E \mid Y = y_i) P(Y = y_i)$$

$$P(X = x) = \sum_{i=1}^{k} P(X = x \cap Y = y_i) = \sum_{i=1}^{k} P(X = x \mid Y = y_i) P(Y = y_i)$$

Law of Total Probability – the details

<u>Theorem</u>: **[Law of Total Probability for Discrete R.V.s]**
Let E be an event. Let Y be a discrete r.v.

$$\boldsymbol{P}\{E\} = \sum_{y} \boldsymbol{P}\{E \cap Y = y\} = \sum_{y} \boldsymbol{P}\{E \mid Y = y\} \cdot \boldsymbol{P}\{Y = y\}$$

For a discrete r.v. X :

$$\boldsymbol{P}\{X = k\} = \sum_{y} \boldsymbol{P}\{X = k \cap Y = y\} = \sum_{y} \boldsymbol{P}\{X = k \mid Y = y\} \cdot \boldsymbol{P}\{Y = y\}$$

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.

$Z' = X + X = 2X$

$(a+b)^n = \sum_{j=0}^{n} \binom{n}{j} a^j b^{n-j}$

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

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

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

X, Y indep.

Proof

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

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

$j > z \Rightarrow z-j < 0$

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

$$= \sum_{j=0}^{z} e^{-\lambda_1}\frac{\lambda_1^j}{j!}\, e^{-\lambda_2}\frac{\lambda_2^{z-j}}{(z-j)!} = \frac{e^{-(\lambda_1+\lambda_2)}}{z!}\sum_{j=0}^{z}\frac{z!}{j!\,(z-j)!}\lambda_1^j\lambda_2^{z-j}$$

$$\binom{z}{j} = \frac{z!}{j!\,(z-j)!}$$

$$= \frac{e^{-(\lambda_1+\lambda_2)}}{z!}\sum_{j=0}^{z}\binom{z}{j}\lambda_1^j\lambda_2^{z-j} \overset{\text{BT}}{=} \frac{e^{-(\lambda_1+\lambda_2)}}{z!}(\lambda_1+\lambda_2)^z$$

Proof. Written out

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

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

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

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

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

Poisson Random Variables - Summary

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

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

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 Poi(λ)
- 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?

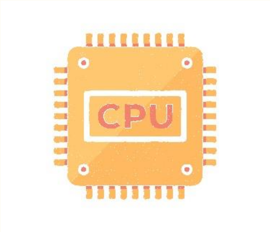
independently

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.

X_1 : # days till disk fails $\sim \text{Geom}(p_1)$

X_2 : # days till CPU fails $\sim \text{Geom}(p_2)$

$\boldsymbol{p_1}$



$\boldsymbol{p_2}$

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

$$P(X_1 < X_2)$$

$$P(E) = \sum_i^n P(E \mid F_i) P(F_i) = \sum^n P(E \cap F_i)$$

$$P(X_1 < X_2 \mid X_1 = k) = \frac{P(X_1 < X_2, X_1 = k)}{\ldots}$$

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 $\sim Geometric(p_1)$

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

$X_1 \perp X_2$

$\boldsymbol{P}\{X_1 < X_2\} =$ $\sum_{k=1}^{\infty} P(X_1 < X_2 \mid X_1 = k) P(X_1 = k)$

$= \sum_{k=1}^{\infty} P(k < X_2) P(X_1 = k)$

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

$P(X_1 < X_2 \text{ and } X_1 = k)$

$= \frac{P(k < X_2, X_1 = k)}{P(X_1 = k)}$

$P(X_2 > k)$

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 $\sim Geometric(p_1)$ X_2 = days until CPU fails $\sim Geometric(p_2)$

$X_1 \perp X_2$

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

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

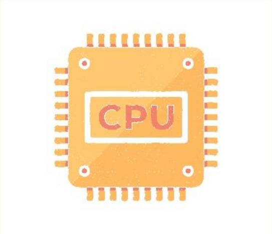
$$= \sum_{k=1}^{\infty} \boldsymbol{P}\{X_2 > k\} \cdot \boldsymbol{P}\{X_1 = k\}$$

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

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.

$\boldsymbol{p_1}$



$\boldsymbol{p_2}$

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

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.

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

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?

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

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.

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

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?

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

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.

components

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

X_i - # days till component i fails $\sim Geo(p)$

D : # days till first component fails

$D = \min(X_1, X_2, \ldots, X_m)$ $\Omega_D = \{1, 2, \ldots$

$P(D = d)$

$P(\underline{D > d}) = P(\underline{\min(X_1, X_2, \dots, X_m) > d})$

$= P(X_1 > d, X_2 > d, \dots, X_m > d)$

$= \underbrace{P(X_1 > d)}_{(1-p)^d} P(X_2 > d) \cdots P(X_m > d)$

$(1-p)^d \cdot (1-p)^d \cdot (1-p)^d$

$= \left[(1-p)^d\right]^m = \left[(1-p)^m\right]^d$

$D \sim \text{Geom}\ \ 1 - (1-p)^m$

$P(Z > d) = \alpha^d$

$Z \sim \text{Geom}\ 1 - \alpha$

End of material for midterm!!!

Next time: continuous random variables!