MATH 180A (Lecture A00)

mathweb.ucsd.edu/~ynemish/teaching/180a

Today: Poisson approximation

Next: ASV 5.1

Week 8:

- Midterm 2 on Wednesday, March 1 (lectures 8-18)
- Homework 5 due Friday, March 3

CLT for Bernoulli / Binomial

Let $S_n \sim \mathrm{Bin}(n,p)$, $a<b$. Then

$$\lim_{n\to\infty} P\left(a \le \frac{S_n - np}{\sqrt{np(1-p)}} \le b\right) = \Phi(b) - \Phi(a)$$

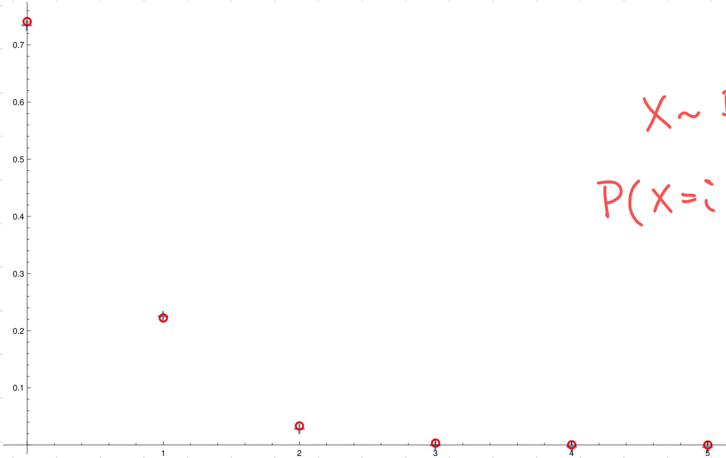
- n independent trials
- probability of success p
- rule of thumb: $np(1-p) > 10$

If this holds, then $N(np, np(1-p))$ gives a good approximation of $\mathrm{Bin}(n,p)$

What about other regimes? $np(1-p) < 10$

Example

$S_{10} \sim \mathrm{Bin}(10, 0.03)$

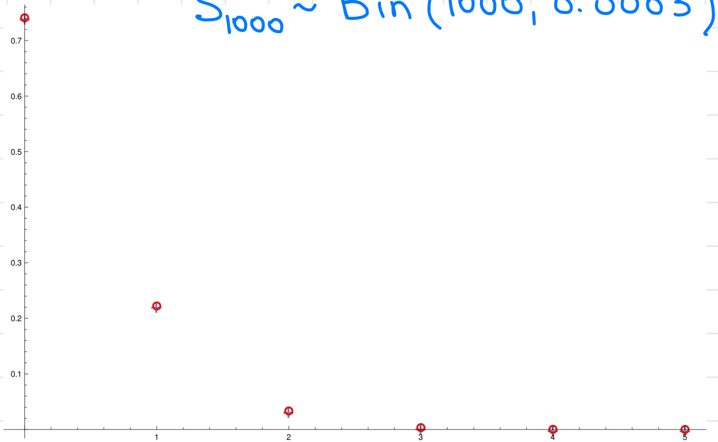


$P(S_{10} = i)$

$X \sim \mathrm{Poisson}(0.3)$

$P(X = i)$

$S_{1000} \sim \mathrm{Bin}(1000, 0.0003)$



$P(S_{1000} = i)$

Poisson approximation

$\lambda > 0$, $X \sim \text{Poisson}(\lambda)$, for $k = 0, 1, 2, \ldots$

$$P(X = k) =$$

Lecture 13: $E(X) =$

$Var(X) =$ (similarly as $E(X)$)

Lecture 12: Let $\lambda > 0$ and n positive integer.

Let $S_n \sim \text{Bin}(n, \frac{\lambda}{n})$ (assume $\frac{\lambda}{n} < 1$).

Then for any fixed $k \in \{0, 1, 2, \ldots\}$

Poisson approximation for finite n

When to approximate by Poisson ?

<u>Proposition</u> Let $X \sim \mathrm{Bin}(n,p)$ and $Y \sim \mathrm{Poisson}(np)$.
Then for any $A \subset \{0,1,2,\dots\}$

If np^2 is small $\to$ use Poisson approximation.

<u>Remark.</u> If $p = \frac{\lambda}{n}$, then one can have $np(1-p) = \lambda(1-\frac{\lambda}{n})$
and at the same time $np^2 = \frac{\lambda^2}{n}$

Then both approximations, normal by $N(\lambda, \lambda)$ and Poisson (λ), are close to $\mathrm{Bin}(n, \frac{\lambda}{n})$

Example. Approximating Bin(n,p)

John flips a coin repeatedly until he tails comes up and counts the number of flips.

Approximate the probability that in a year there are at least 3 days when he needs more that 10 flips.

Approximating probabilities of rare events

Poisson distribution is used to model the occurances of rare events. Examples:

customers arriving in a store ⟷ all potential customers decide independently to come or not

number of emergency calls ⟷ all people in the city have an emergency or not "independently" of each other

number of car accidents ⟷ all drivers in the county have accidents (or not) "independently"

number of goals scored in a hockey game ⟷ a lot of "independent" shots

Example

Number of phone calls in a day can be modeled by Poisson random variable. We know that on average 0.5% of the time the call center receives no calls at all. What is the average number of calls per day?

Exercise

10% of households earn > 80000 $

0.25% of households earn > 450000 $

Choose 400 households at random. Denote

$X = \#$ households > 80000 , $Y = \#$ households > 450000 $

Estimate $P(X \geq 48)$ and $P(Y \geq 2)$

Question

A fair 20-sided die is tossed 400 times.
We want to calculate the probability that a 13 came up at least 25 times. We should use

(a) Poisson approximation
(b) Normal approximation
(c) Neither
(d) Both

Waiting for a customer

Suppose that customers arrive in a store withe rate λ customers per hour. How can we model the time until the first (or next) customer arrives?

Additional assumptions: if the intervals are small enough, then

- only one customer can arrive per interval
- customers arrive/do not arrive for each interval independently
-

Q: What is $P(X>t)$

Exponential distribution

$P(X>t)$

CDF: $P(X \le t)$

PDF: $f_X(t) =$

<u>Def</u>. Let $\lambda > 0$. We say that random variable X has exponential distribution with rate parameter λ, if

$$f_X(t) = \begin{cases} \lambda e^{-\lambda t}, & t > 0 \\ 0, & t \le 0 \end{cases}, \quad \text{denote } X \sim \mathrm{Exp}(\lambda)$$

Exponential distribution

Let $X \sim Exp(\lambda)$. Then

- $E(X) = \int_0^\infty t\lambda e^{-\lambda t} dt = \int_0^\infty t\left(-e^{-\lambda t}\right)' dt = -t\lambda e^{-\lambda t}\Big|_0^\infty + \int_0^\infty e^{-\lambda t} dt = \frac{1}{\lambda}$
- $E(X^2) = \frac{2}{\lambda^2}$, so $Var(X) = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$
- $P(X > t) = e^{-\lambda t}$

Exponential distribution is used to model waiting times

<u>Example</u> Length of a phone call is modeled by exponential random variable with mean 10 (minutes). What is the probability that the call takes > 8 minutes? Between 8 and 22?

Memoryless property

Proposition Let $X \sim Exp(\lambda)$, $\lambda > 0$. Then for any $s, t > 0$

Proof $P(X > s+t \mid X > s)$

$Exp(\lambda)$ is the only continuous distribution with memoryless property.

Remark. If $N \sim Geom(p)$, then $P(N > k) = (1-p)^k$, and

$$P(N > k+l \mid N > k) = \frac{P(N > k+l)}{P(N > k)} = \frac{(1-p)^{k+l}}{(1-p)^k} = (1-p)^l = P(N > l)$$

Example

Animals are crossing a highway. Intervals between arriving cars have exponential distribution with mean 30 (min). Turtle needs 10 minutes to cross.

(a) What is the probability that the turtle survives?

(b) When the turtle starts crossing the highway, a racoon says that it has not seen a car for 5 minutes. Will this change the survival probability?

New section

Characterizing random variables

- PMF / PDF for discrete / continuous random variables

$$P(X \in A) = \sum_{t \in A} p_X(t), \quad P(X \in A) = \int_A f_X(t)\,dt$$

- CDF

$$F_X(t) = P(X \leq t)$$

- $E(X)$, $Var(X)$ gives partial information

- Moments $\left(E(X^k)\right)_{k \geq 1}$ (sometimes) describe uniquely the distribution

NEW TOOL: Moment generating function (MGF)

convenient when working with sums of independent RVs.

Moment generating function

__Def__. Let X be a random variable. Then

__Examples__ (more in the textbook)

- $X \sim Ber(p)$, $E(e^{tX}) =$
- $X \sim Poisson(\lambda)$, $E(e^{tX}) =$
- $X \sim N(0,1)$, $E(e^{tX}) =$

Moment generating function

Examples

- $X \sim N(\mu, \sigma^2)$, use that if $Z \sim N(0,1)$, then $\sigma Z + \mu \sim N(\mu, \sigma^2)$

  $E(e^{tX}) =$

- Exercise :

  $X \sim Exp(\lambda)$

  $M_X(t) =$