DAV-2

Problem Solving - 2

→ Distributions
→ Questions
→ Log Normal Distribution.

Tuesday's class.

Distribution	Parameters	Equation	Mean	Variance	Code
Bernoulli	p	$P[X=1]=p$ $P[X=0]=1-p$	p	$p(1-p)$	bernoulli.pmf()
Binomial	n, p	$P[X=k]={}^{n}C_{k}p^{k}(1-p)^{n-k}$	np	$np(1-p)$	binom.pmf()
Geometric	p	$P[X=k]=(1-p)^{k-1}p$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	geom.pmf()
Poisson	λ	$P[X=k]=\frac{\lambda^k e^{-\lambda}}{k!}$	λ	λ	poisson.pmf()
Exponential	λ	$F(x)=1-e^{-\lambda x}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	expon.cdf()
Normal	μ, σ	$f(x)=\frac{1}{2\sigma^2}e^{\frac{(x-\mu)^2}{\sigma^2}}$	μ	σ^2	norm.cdf()
Log normal	μ, σ	$\log X \sim$ Gaussian	$exp(\mu+\sigma^2/2)$	$e^{(\sigma^2)-1}e^{2\mu+\sigma^2}$	lognorm.cdf()

Retail supply chain. (Toothpaste) ??

S3 S1 S2 S4

DC

Manage Inventory
(Optimizes on replenishment)

↓

Every Sunday.

S1

Mean.

Q Weekly Sales → 1000 toothpaste/week.

 → 1300 slots

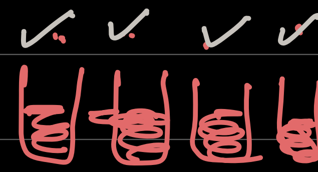
$$P(X>1300) = 1 - P(X \leq 1300)$$

$$= 1 - \text{poisson.cdf}(k = 1300, \ mu = 1000)$$

Poisson ✓

Binomial ✗

Q Pooled Blood Test: optimise # tests

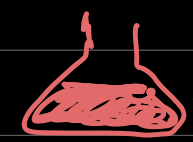


→ Pool samples of 4 people at a time

→ If clean, all 4 samples are clean.

→ If bad, 4 samples are tested individually

→ P(unacceptable) = 0.1



Pool

Q Find the expected number of tests - <4 $=4$ >4

Pool ✓ ✗ ✗

$X \to$ # tests. $E[X]$??? $p = 0.1$

Thm of total prob.

$0.1 \times \cancel{4} + \cancel{0.9} \times 1$ — $P[X=1] = \boxed{{}^4C_0 (1-p)^4 p^0} \times 1$

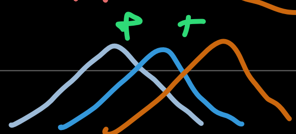
5

$P[X=5] = 1 - P[X=1] \times 5$ **

Battery (lifetime)   Mean lifetime $= 5$ weeks
std $= 1.5$ weeks

Approx. 13 or more batteries will be needed in a year ???

in a year $\rightarrow$ 52 weeks.



$\Downarrow$

Total cum. lifetime of 12 batteries $<$ 52 weeks

$X_i \rightarrow$ lifetime of ith battery

$Y \rightarrow X_1 + X_2 + X_3 \ldots + X_{12} \rightarrow$ Total lifetime of 12 battery.

$E[Y] = E[X_1] + E[X_2] + E[X_3] \ldots = 12\,E[X_i] = 12 \times 5 = 60$

$Var[Y] = Var[X_1] + Var[X_2] + \ldots = 12\,Var[X] = 12 \times (1.5)^2 = 27$

$Std[Y] = \sqrt{27}$

$$P[Y < 52] = P\left[Z < \frac{52 - 60}{\sqrt{27}}\right] = P\left[Z < \frac{-8}{\sqrt{27}}\right]$$

cdf

Airline Overbooking

5% will not show up

Capacity → 50 seats

Suppose 52 tickets → Overbooking

What is P that everyone who turns gets a seat?

$p \rightarrow$

$n \rightarrow 52$

$X \rightarrow$ # passenger who show up.

$X \leq 50$

$P[X \leq 50] = \text{binom.cdf}(k = 50,\ n = 52,\ p = 0.95)$ ⊕ Earning

$P[51] \rightarrow$ cost

$P[52] \rightarrow$ cost

$\Big\}$ −cost

Hiring

IQ is known to be approximately Gaussian with a mean of 100 and standard deviation of 15

You want to hire a person with IQ greater than 110

Conducting an interview costs 1000. What is the expected budget?

$$P[X > 110] = P\left[Z > \frac{(110 - 100)}{15}\right] = \text{1 - norm.cdf(10/15)} = 0.252$$

Let N denote the number of interviews till first hire

Which distribution does N follow? Geometric

What is $E[N]$? 1/0.252 = 3.96

Expected budget = 3960

Simulate a fair coin from a biased coin

There is a coin that lands heads 70% of the times

How can we use this coin so that it lands heads 50% of the times?

Verify the output of your algorithm using 10000 simulations at 95% confidence

Let us toss the biased coin twice

Sample space S = {HH, HT, TH, TT}

$$P[HH] = 0.7 * 0.7$$
$$P[HT] = 0.7 * 0.3$$
$$P[TH] = 0.3 * 0.7$$
$$P[TT] = 0.3 * 0.3$$

These two have same probability

Biased	Fair
$HH \rightarrow$	Ignore
$TT \rightarrow$	Ignore
$HT \rightarrow$	H
$TH \rightarrow$	T