

Lab 2

Probability review

Pairwise independence does not imply mutual independence.

- $X = \text{Flip coin, } \Pr(X = \text{heads}) = \frac{1}{2}$
- $Y = \text{Flip Coin, } \Pr(Y = \text{heads}) = \frac{1}{2}$
- $Z = \Pr(\text{exactly } X \text{ or } Y \text{ are heads (not both)}) = \frac{1}{2}$

X	Y	Z = XOR(X,Y)
0	0	0
1	0	1
0	1	1
1	1	0

X	Y	Pr[X,Y]
0	0	1/4
1	0	1/4
0	1	1/4
1	1	1/4

X	Z	Pr[X,Y]
0	0	1/4
1	1	1/4
0	1	1/4
1	0	1/4

Pairwise independence does not imply mutual independence.

- $\Pr(X,Z)$ pairwise independence:
 - $\Pr(X=0,Z=0) = \Pr(X=0) * \Pr(Z=0) = \frac{1}{2} * \frac{1}{2} = \frac{1}{4}$
- NOT mutual independence
 - $\Pr(X=1,Y=1,Z=0) = \frac{1}{4}$ where as $\Pr(X=1)*\Pr(Y=1)*\Pr(Z=0) = 1/8$

X	Y	Z = XOR(X,Y)	$\Pr[X]*\Pr[Y]*\Pr[Z]$	$\Pr[X,Y,Z]$
0	0	0	$\frac{1}{8}$	$\frac{1}{4}$
1	0	1	$\frac{1}{8}$	$\frac{1}{4}$
0	1	1	$\frac{1}{8}$	$\frac{1}{4}$
1	1	0	$\frac{1}{8}$	$\frac{1}{4}$
0	0	1	$\frac{1}{8}$	0
1	0	0	$\frac{1}{8}$	0
0	1	0	$\frac{1}{8}$	0
1	1	1	$\frac{1}{8}$	0

Definitions

- Expectation

- When X is a discrete random variable

- $\sum_i i \cdot \Pr[X = i]$

- When X is a continuous random variable, and $f(x)$ is its probability density function

- $\int_{-\infty}^{\infty} x \cdot f(x) dx$

- Variance

- $Var[X] = E[(X - E[X])^2] = E[X^2] - E[X]^2$

- It measures how far your data spread out from their average value.

- Standard deviation

- $\sigma_X = \sqrt{Var[X]}$

Covariance and Correlation

Definitions: $Cov(X, Y) = E[(X - E[X])(Y - E[Y])]$

$$= \sum_x \sum_y (x - E[X])(y - E[Y])p_{X,Y}(x, y)$$

$$Corr(X, Y) = \frac{Cov(X, Y)}{\sigma_X \sigma_Y}$$

From a linear algebra perspective: If X, Y are two random variables of zero mean, then the covariance $Cov[XY] = E[X \cdot Y]$ is the dot product of X and Y . The standard deviation of X is the length of X . The correlation is the cosine of the angle between the two vectors.

Covariance

Some properties:

$$\text{Cov}(X + c, Y) = \text{Cov}(X, Y)$$

$$\text{Cov}(aX + bY, Z) = a \cdot \text{Cov}(X, Z) + b \cdot \text{Cov}(Y, Z)$$

Exercise:

X and Y are two independent $\mathcal{N}(0, 1)$ random variables and:

$$Z = 1 + X + XY^2$$

$$W = 1 + X$$

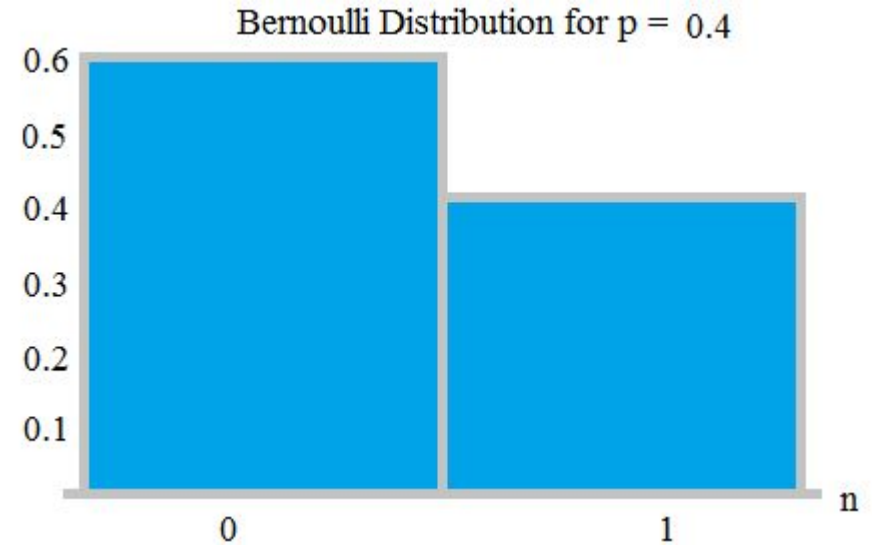
Find $\text{Cov}(W, Z)$

Discrete probability distributions

	PMF	Mean	Variance
Bernoulli (p)	1-p if k=0 p if k=1	p	p(1-p)
Binomial (n,p)	$\binom{n}{k} p^k (1-p)^{n-k}$	np	np(1-p)
Poisson (λ)	$\frac{\lambda^k e^{-\lambda}}{k!}$	λ	λ
Geometric (p)	$(1-p)^{k-1} p$	$\frac{1}{p}$	$\frac{1-p}{p^2}$

Bernoulli(p)

- Discrete
- Two outcomes: 0,1 | | fail, success
- $\Pr(X = 1) = p = 1-q$
- $\Pr(X = 0) = q = 1-p$
- Expectation:
 - $E[x] = 1 \cdot p + (0) \cdot (1-p) = p$



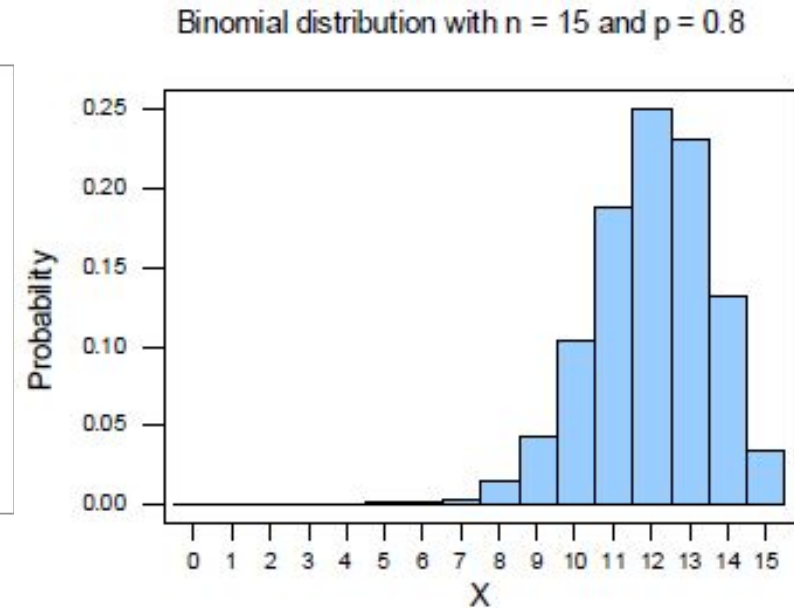
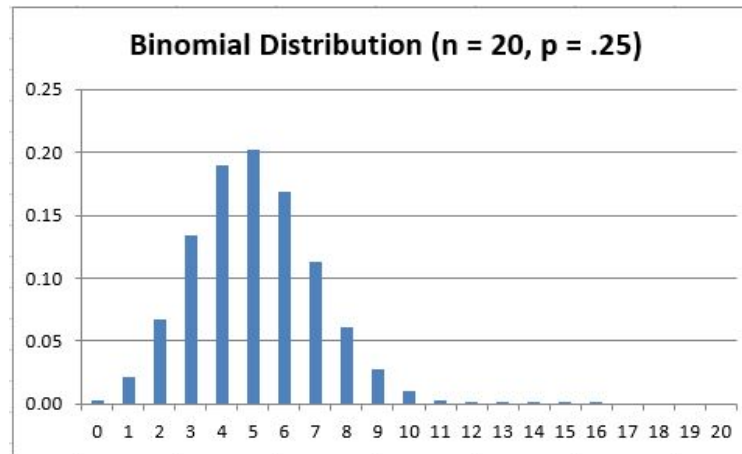
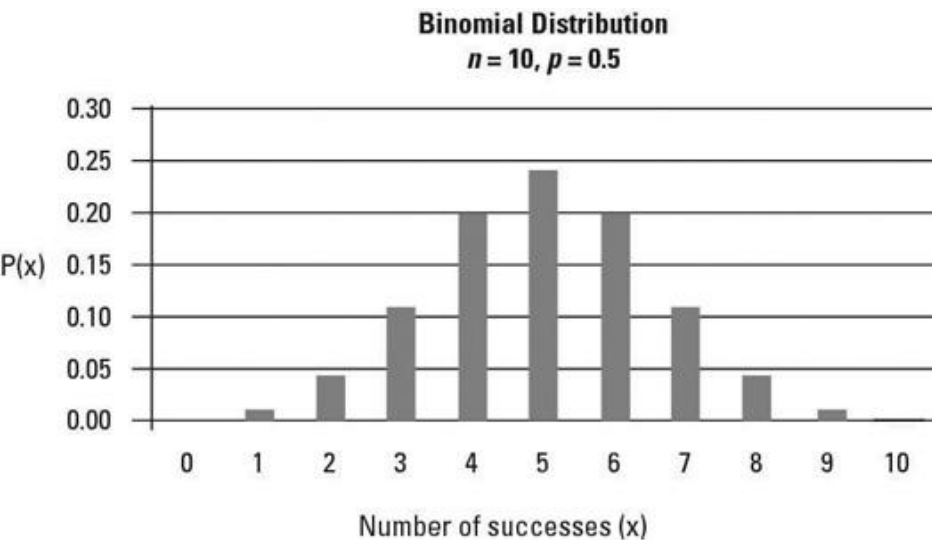
Binomial(n,p)

Example: Flip a coin 5 times, what is the probability of having exactly 2 heads?

- Sample repetitively from a Bernoulli distribution n times
- Bernoulli = Binomial with n=1
- $\Pr[X = k] = \binom{n}{k} p^k (1 - p)^{n-k}$
- Expectation
 - Use linearity of expectations
 - Let X_i be the ith sample outcome from the Bernoulli distribution
 - $E[X_1 + \dots X_n] = \sum_i E[X_i] = n \cdot p$

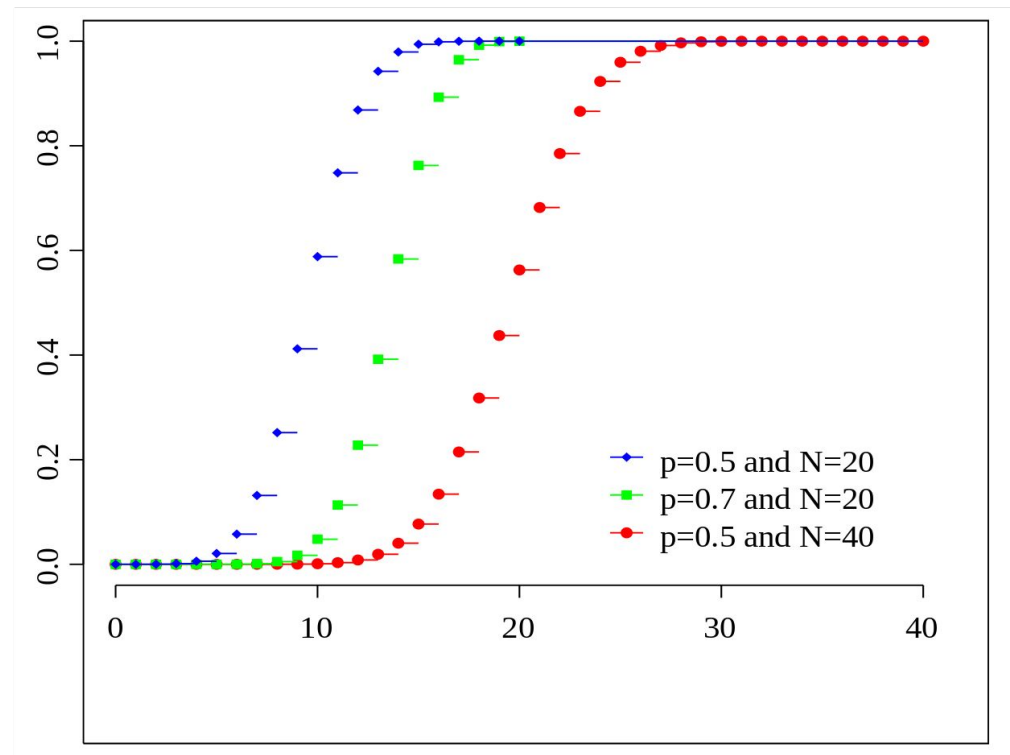
Binomial Distributions

PDF $\Pr[X = k] = \binom{n}{k} p^k (1 - p)^{n-k}$



Binomial Distribution CDF

- **CDF:** $F_x(t) = \Pr(X \leq T)$



Poisson(λ)

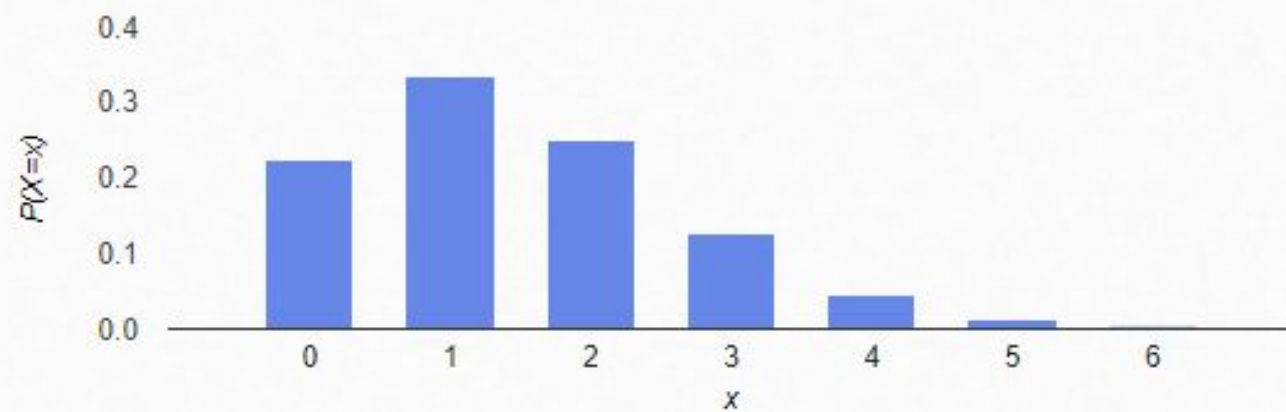
Poisson distribution is a **probability distribution that is used to show how many times an event is likely to occur over a specified period**

- Events are independent of each other.
- The occurrence of one event does not affect the probability another event will occur.
- The average rate (events per time period) is constant.
- Two events cannot occur at the same time.

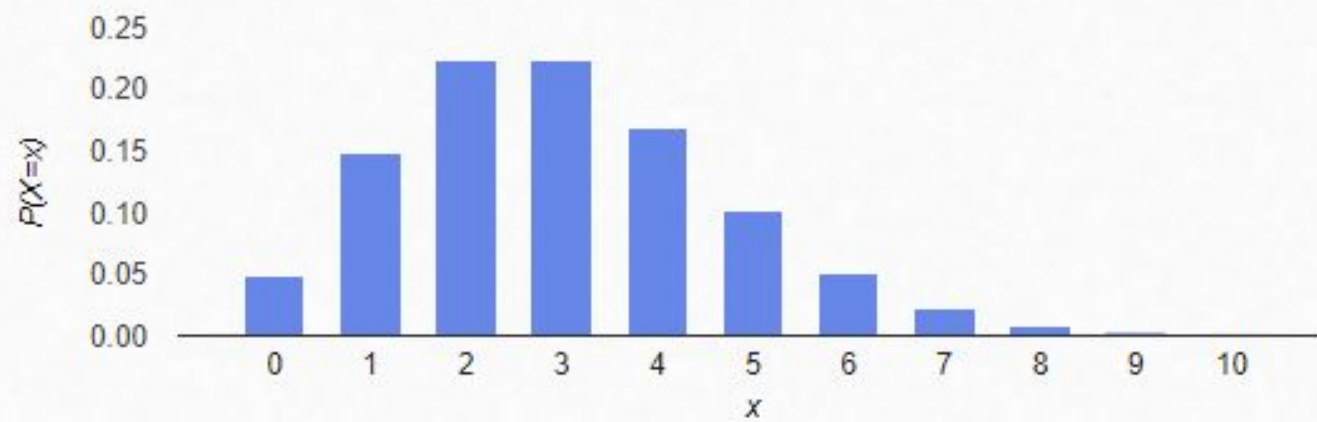
$$\Pr[X = k] = \frac{\lambda^k e^{-\lambda}}{k!}$$

Poisson Example

- Example: You are “on call” at work. And every night between 12 and 8 AM you receive about 6 calls from customers
 - Question: What is the probability you receive a call from 2-4 AM?
 - Definitions: $X \sim P(\lambda)$ where X is a random variable with a Poisson distribution, and “ λ ” is the mean for the interval
 - 6 calls over an 8 hour window is about $\frac{3}{4}$ of a call every hour, our window is 2 hours, so our “ λ ” value is $2 * \frac{3}{4} = 1.5$
- <https://homepage.divms.uiowa.edu/~mbognar/applets/pois.html>

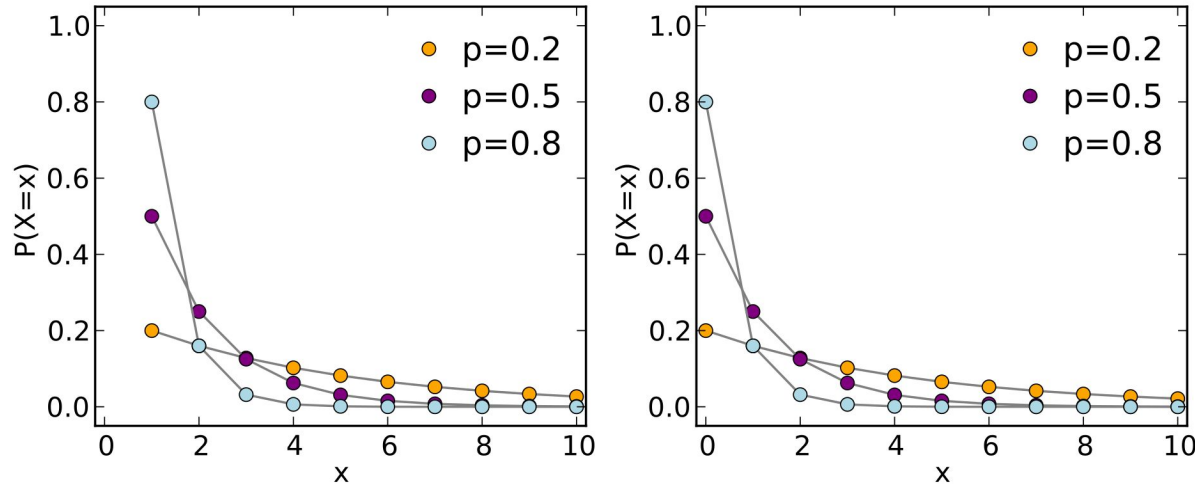


$$\mu = E(X) = 1.5 \quad \sigma = SD(X) = 1.225 \quad \sigma^2 = Var(X) = 1.5$$



$$\mu = E(X) = 3 \quad \sigma = SD(X) = 1.732 \quad \sigma^2 = Var(X) = 3$$

Geometric(p)



Expected Value:

$$\sum_{k=1}^{\infty} k * (1 - p)^{k-1} p$$

- Discrete, closely related to Bernoulli
- Probability of N failures before a success
- $Pr(X = k) = (1 - p)^{k-1} p^1$
- Example: What is the probability I flip a coin 4 times before getting heads?

Geometric Expectation Derivation

$$1. \quad E[X] = \sum_{k=1}^{\infty} k * (1 - p)^{k-1} p$$

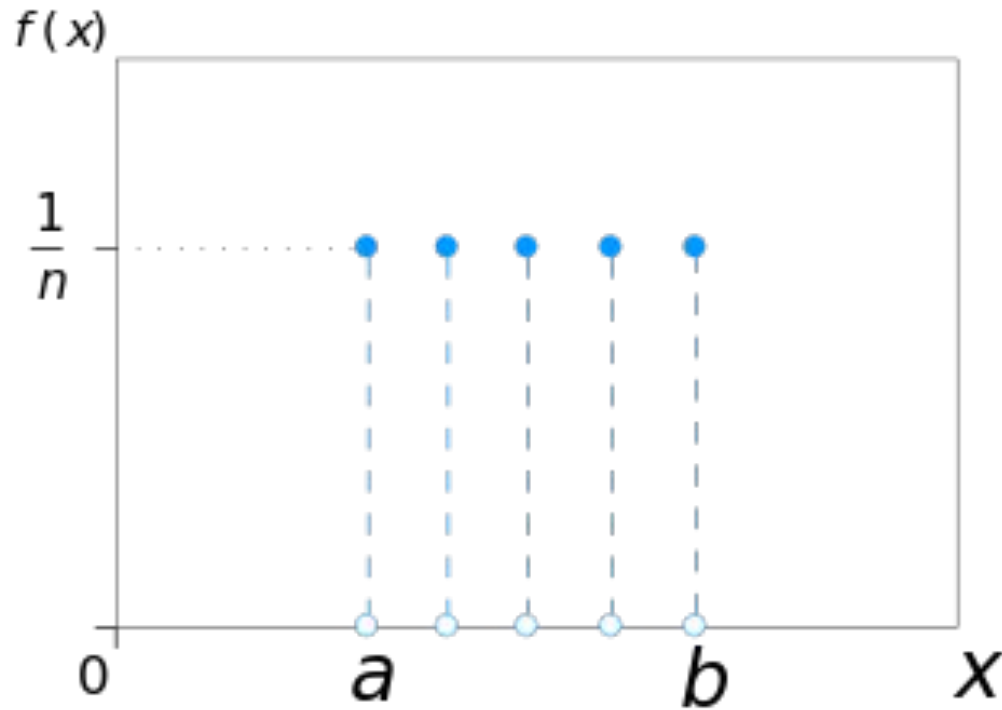
$$2. \quad = p * \sum_{k=1}^{\infty} k * (1 - p)^{k-1} \quad // \text{ take out } p \text{ from summation}$$

$$3. \quad = p * \sum_{k=1}^{\infty} \frac{d}{dp} * -(1 - p)^k \quad // \text{ substitute derivative w.r.t } p$$

$$4. \quad = p * -\frac{d}{dp} \sum_{k=1}^{\infty} (1 - p)^k \quad // \text{ well known series summation}$$

$$5. \quad = -p * \frac{d}{dp} * \frac{1}{p} = -p \left(-\frac{1}{p^2} \right) = \frac{1}{p}$$

Uniform - Discrete



$E[X]$ where $P(X=x) = 1/N$:

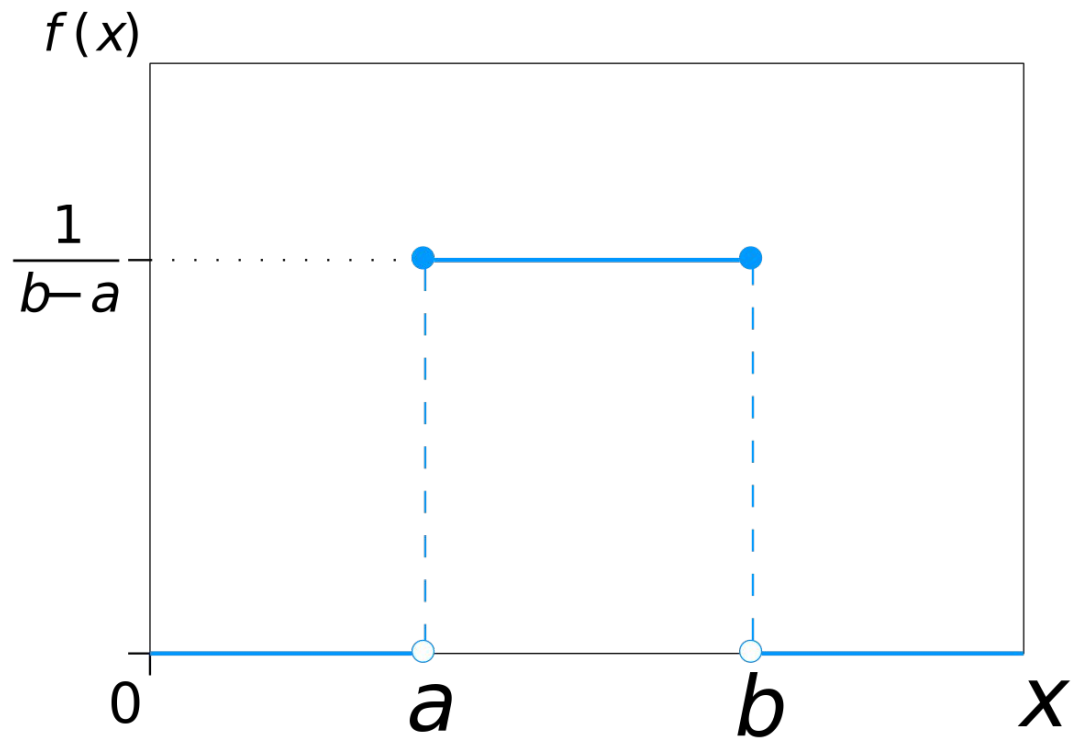
$$(N+1)/2$$

$$\text{Var}[x] = \frac{N^2 - 1}{12}$$

Continuous probability distributions

	PDF	Mean	Variance
Uniform (a,b)	$\frac{1}{b-a} \quad \text{for } a \leq x \leq b$ 0 otherwise	$(a+b)/2$	$(b-a)^2/12$
Exponential (λ)	$\lambda e^{-\lambda x}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Laplace (μ, b)	$\frac{1}{2b} \exp\left(-\frac{ x-\mu }{b}\right)$	μ	$2b^2$
Gaussian (μ, σ^2)	$\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$	μ	σ^2

Uniform - Continuous

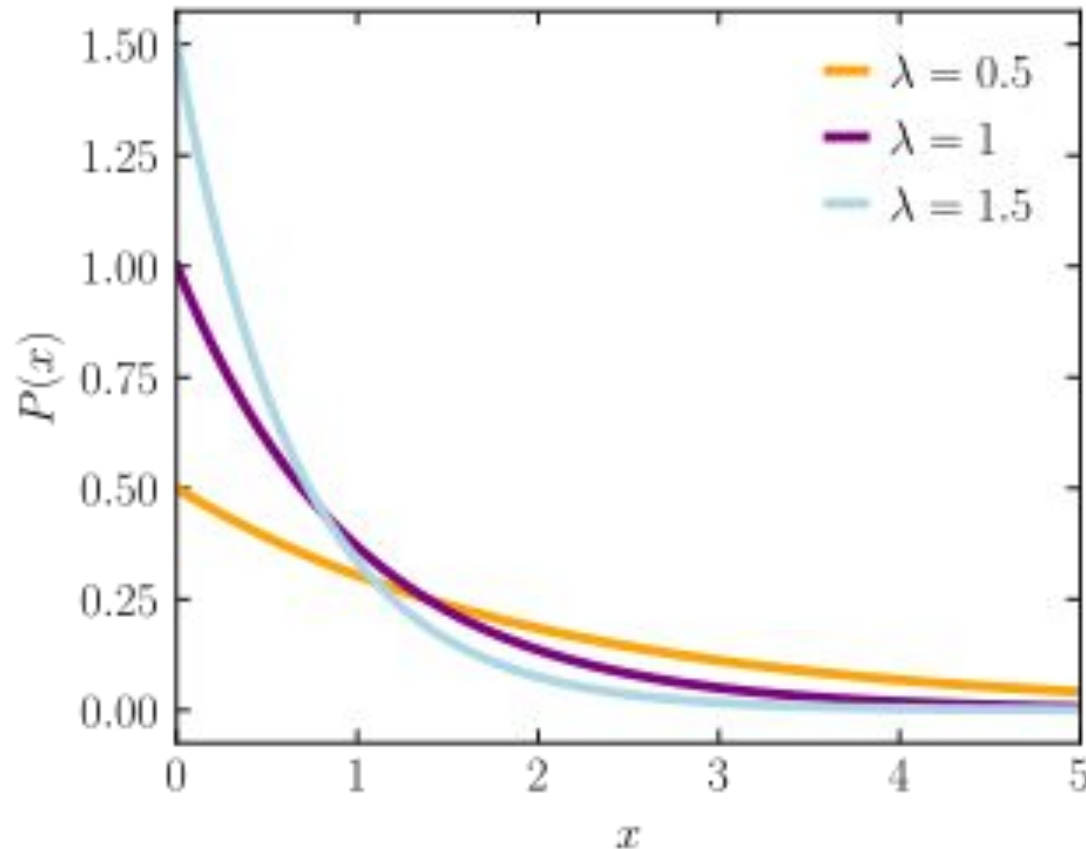


$E[X]$ where $P(X=x) = 1/N$:

$$\frac{1}{b-a}$$

$$\text{Var}[x] = \frac{(b-a)^2}{12}$$

Exponential

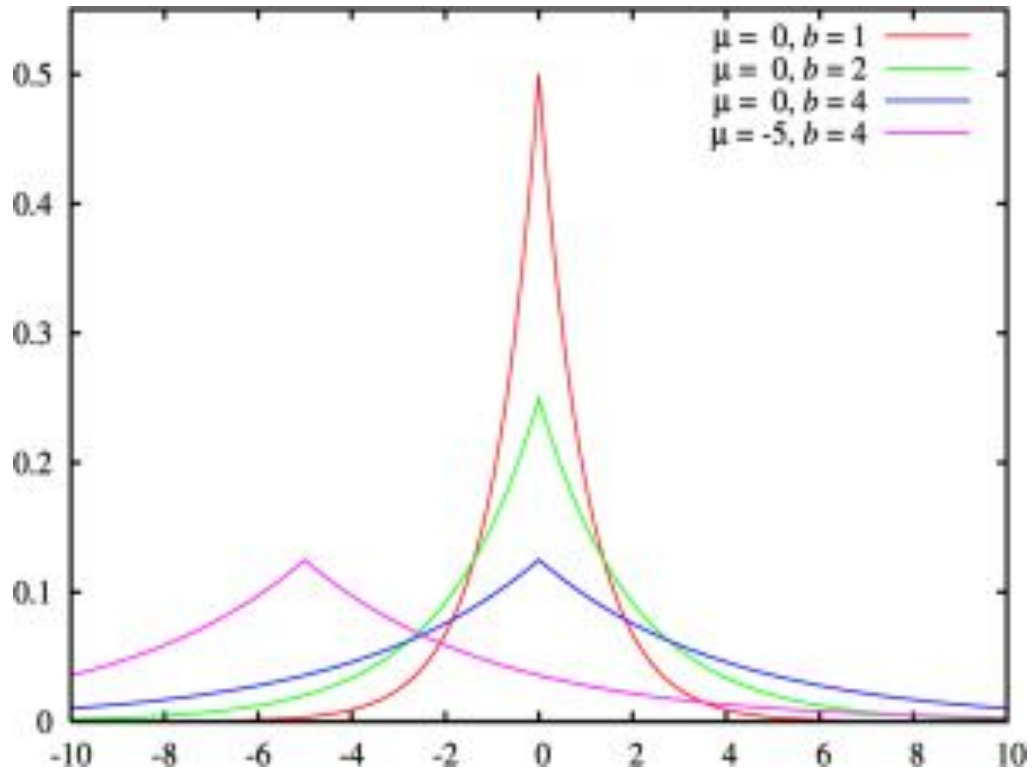


The exponential distribution is the probability distribution of the time between events in a Poisson point process.

- Parameterized by (λ)
- PDF $\Pr[X = x] = \lambda e^{-\lambda x}$
- Expected value $1/\lambda$
- Variance $1/\lambda^2$

Remark: Exponential distribution is not the same as the class of exponential families of distributions

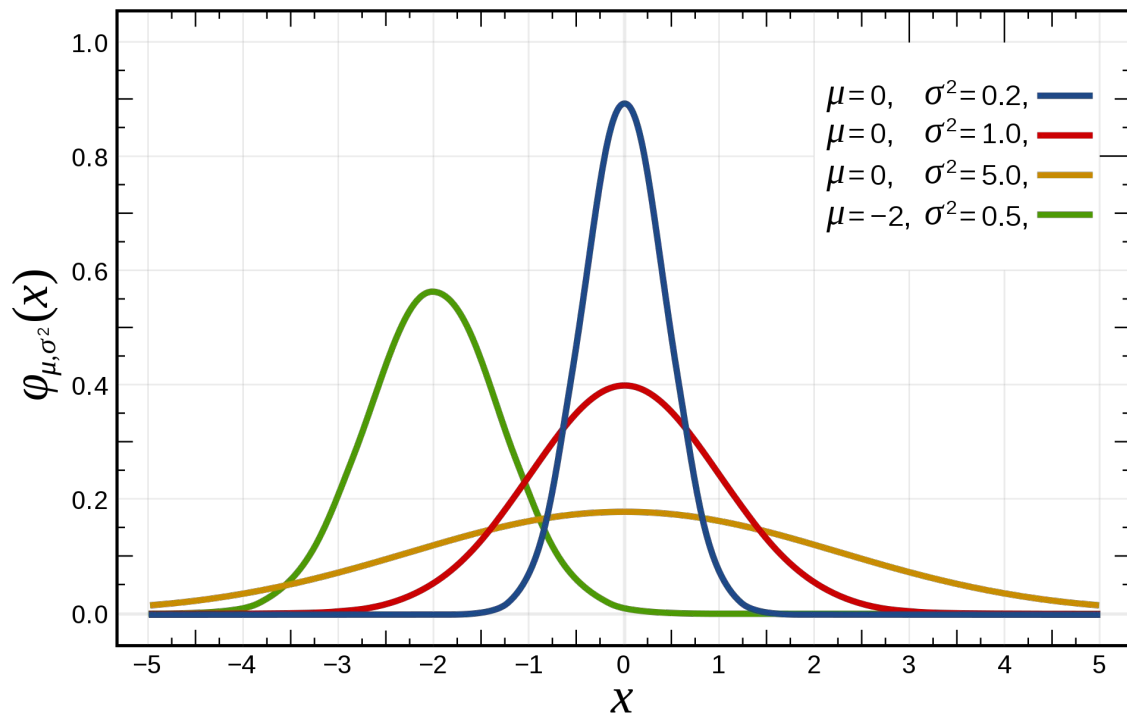
Laplace (μ, b)



- It can be thought of as two exponential distributions (with an additional location parameter) spliced together along the abscissa
- Parametrized by mu and beta:
 - Mu = Location (shift along the x axis)
 - B = scaling, maximum peak value

$$\Pr[X = x] = \frac{1}{2b} \exp \left(-\frac{|x - \mu|}{b} \right)$$

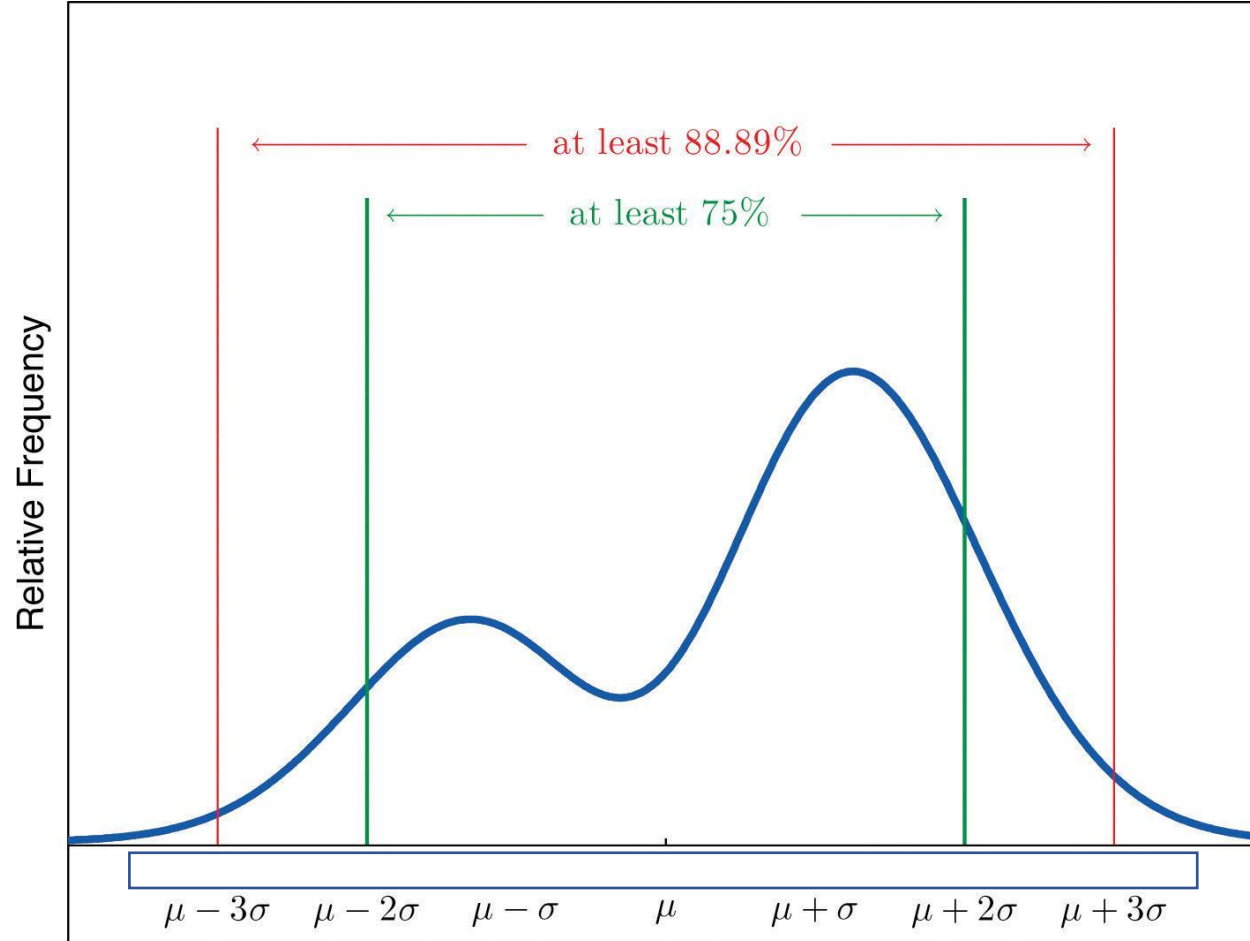
Gaussian – One dimension



$$\Pr[X = x] = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right)$$

- Parameterized by
 - μ = mean
 - σ = standard deviation
- It is important partly due to CLT.

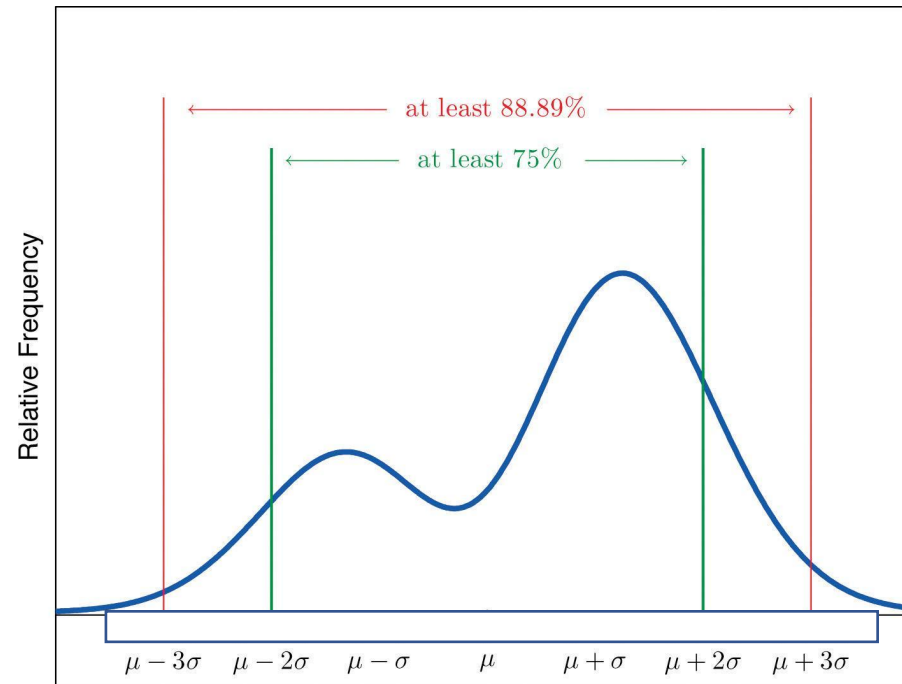
Chebyshev Inequality



- $P(|X - \mu| \geq k * \sigma) \leq \frac{1}{k^2}$

Poisson Chebyshev problem

- Let us say that X is modeled by a Poisson variable with $\mu/\lambda/\text{mean} = 100$
- What is the lower bound of $P(75 < X < 125)$?



Poisson Chebyshev solution

- Find μ and σ :

- $\mu = 100$

- $\sigma = \sqrt{var} = \sqrt{100} = 10$

- $P(|X - \mu| \geq k * \sigma) \leq \frac{1}{k^2}$ is the P that we exist outside a certain SD

- What we want is the P we exist WITHIN a certain SD

- $P(75 < X < 125) = 1 - P(|X - 100| \geq 2.5 * 10) \geq 1 - \frac{1}{2.5^2} = 0.84$

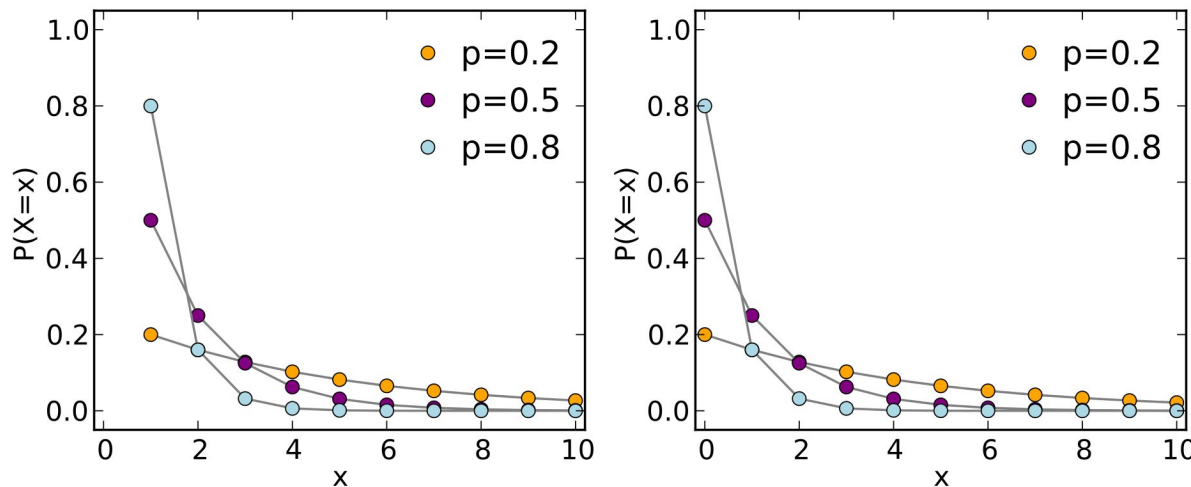
End

Practice Problem

- Broken device, calling customer support
- 1% chance a human representative being available and picking up my call, or else robot will disconnect me
- Question:
 - What is the expected number of times I will have to call in before I talk to a human being?

Practice Solution

- Identify an analytical distribution:
 - Geometric in our case (fails until a success)



- Discrete, closely related to Bernoulli
- Probability of N failures before a success
- $P(X = k) = (1 - p)^{k-1} p^1$
- Example: What is the probability I flip a coin 4 times before getting heads?

Practice Solution

- X = number of call attempts before a human picks up
- $p = .01$, $q = 1-p$
- $E[X] = \sum_{i=0}^N X_i * P(X_i)$
- $= 1(.01) + 2(.99)^1(.01) + 3(.99)^2(.01) + \dots$
- $= p/(1 - q)^2 = 1/p = 1/.01 = \mathbf{100}$