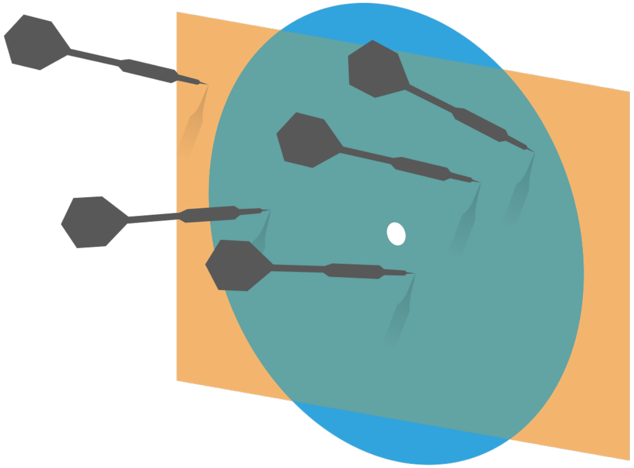
Probability and Statistics for Computer Science



Credit: wikipedia

Who discovered this?

$$e = \lim_{n \to \infty} \left(1 + \frac{1}{n}\right)^n$$

Hongye Liu, Teaching Assistant Prof, CS361, UIUC, 02.25.2020

Last time

- Random Variable
  - *Review* , Covariance
  - *The weak law of large numbers*

Proof of Weak law of large numbers

- Apply Chebyshev's inequality $\forall \varepsilon > 0$

$$P(|\overline{\mathbf{X}} - E[\overline{\mathbf{X}}]| \geq \epsilon) \leq \frac{var[\overline{\mathbf{X}}]}{\epsilon^2}$$

- Substitute $E[\overline{\mathbf{X}}] = E[X]$ and $var[\overline{\mathbf{X}}] = \dfrac{var[X]}{N}$

$$P(|\overline{\mathbf{X}} - E[\mathbf{X}]| \geq \epsilon) \leq \frac{var[\mathbf{X}]}{N\epsilon^2} \xrightarrow[N \to \infty]{} 0$$

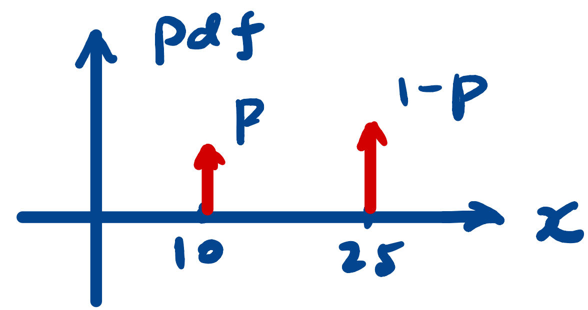
$$\boxed{\lim_{N\to\infty} P(|\overline{\mathbf{X}} - E[X]| \geq \epsilon) = 0}$$

Random Variable Example

Money box

10¢ 25¢

dime quarter



* shake and take one and put back.

$X_1, X_2, \dots X_N$ are iid.

X_1 takes $x_1 = 10$

X_2 takes $x_2 = 10$

X_3 takes $x_3 = 25$

$\vdots$

X_N takes $x_N = 10$

$E[X] = ?$ $10 \cdot p + (1-p) \cdot 25$

$$\bar{X} = \frac{\sum X_i}{N}$$

Applications of the Weak law of large numbers

- The law of large numbers ***justifies using simulations*** (instead of calculation) to estimate the expected values of random variables HW4

$$\lim_{N \to \infty} P(|\overline{\mathbf{X}} - E[X]| \geq \epsilon) = 0$$

- The law of large numbers also ***justifies using histogram*** of large random samples to approximate the probability distribution function $P(x)$, see proof on

Pg. 353 of the textbook by DeGroot, et al.

review on your own

Histogram of large random IID samples approximates the probability distribution

- The law of large numbers justifies using histograms to approximate the probability distribution. Given ***N*** IID random variables X_1, …, X_N
  - According to the law of large numbers

$$\mathbf{\overline{Y}} = \frac{\sum_{i=1}^{N} Y_i}{N} \xrightarrow{N \to \infty} E[Y_i]$$

  - As we know for indicator function textbook pg 100-101

$$E[Y_i] = P(c_1 \leq X_i < c_2) = P(c_1 \leq X < c_2)$$

What is the number?

$$(p+q)^n = \sum_{k=0}^{n} a_k p^k q^{n-k}$$

$a_k = ?$ $\binom{n}{k}$ → Binomial coeff.

if $p+q=1$

$LHS = 1^n = 1$ RH $= 1$

therefore $\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = 1$

How many paths that lead A→B ?

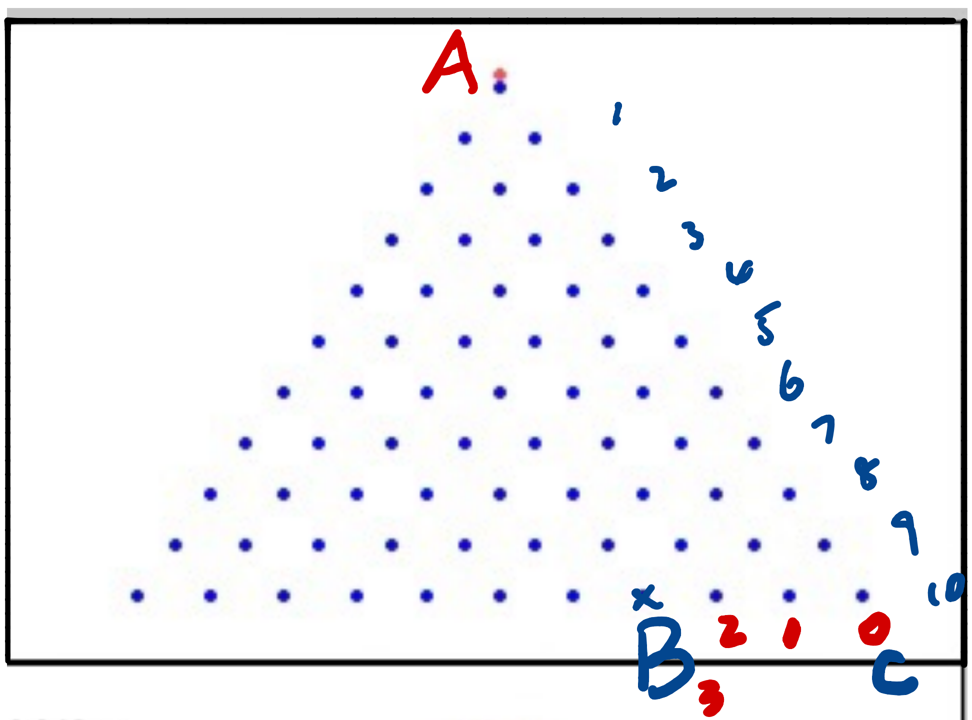
L _ _ L _ _ _ L _ _

R? R?

10

$\binom{10}{3}$ or $\binom{10}{7}$

$\binom{10}{K}$



Objectives

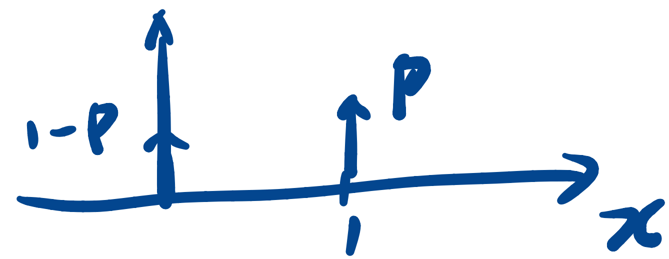
- Important known discrete probability distributions
- Continuous Random Variable

The classic discrete distributions

H T coin

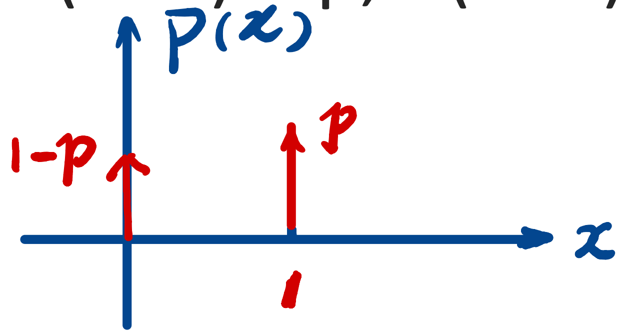
$$X(\omega) = \begin{cases} 1 & H \\ 0 & H^c \end{cases}$$

- Bernoulli
- Binomial sum of N Bernoulli:
- Geometric
- Discrete uniform



Bernoulli distribution

- A random variable X is **Bernoulli** if it takes on two values 0 and 1 such that P(X=1) = p, P(X=0)=1-p



$$E[X] = p$$

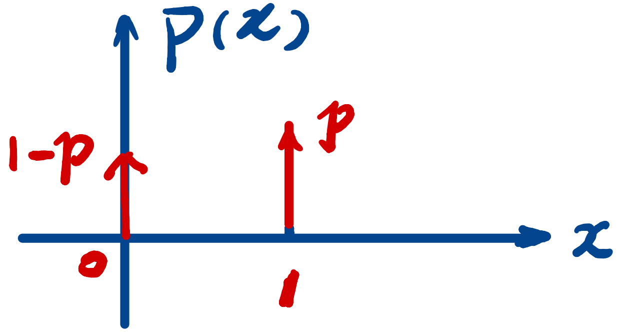
$$var[X] = p(1-p)$$

Jacob Bernoulli (1654-1705)

Credit: wikipedia

Bernoulli distribution

$$E[X] = p \cdot 1 + (1-p) \cdot 0$$
$$= p$$



$$\mathrm{var}[X] = E[X^2] - (E[X])^2$$
$$= 0^2 \cdot (1-p) + 1^2 \cdot p - p^2$$
$$= p - p^2$$
$$= p(1-p)$$

Bernoulli distribution

- Examples
  - Tossing a biased (or fair) coin
  - Making a free throw
  - Rolling a six-sided die and checking if it shows 6
  - **Any indicator function** of a random variable

$$I_A = \begin{cases} 1 & \text{event } A \text{ happens} \\ 0 & \text{otherwise} \end{cases}$$

$$E[I_A] = 1 \cdot P(A) + \cancel{0 \cdot P(A^c)}$$
$$= P(A)$$

Binomial distribution

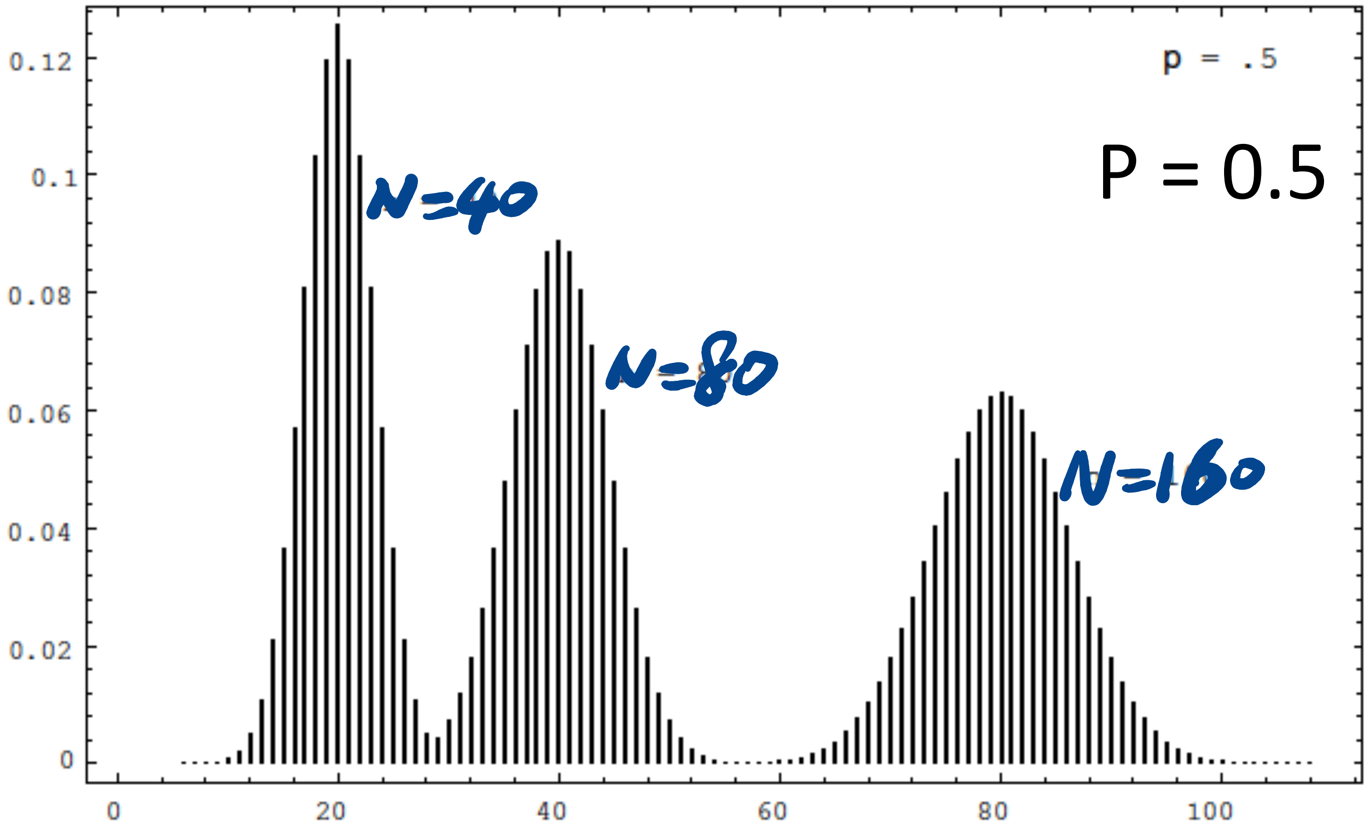
- The Galton Board

  http://www.randomservices.org/random/apps/GaltonBoardExperiment.html

- Remember the airline problem?

Binomial distribution

N



p

Credit: Prof. Grinstead

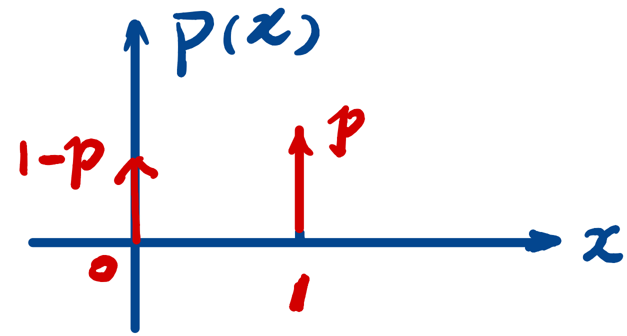
Binomial distribution

Binomial RV X_s is the sum of N independent Bernoulli RVs

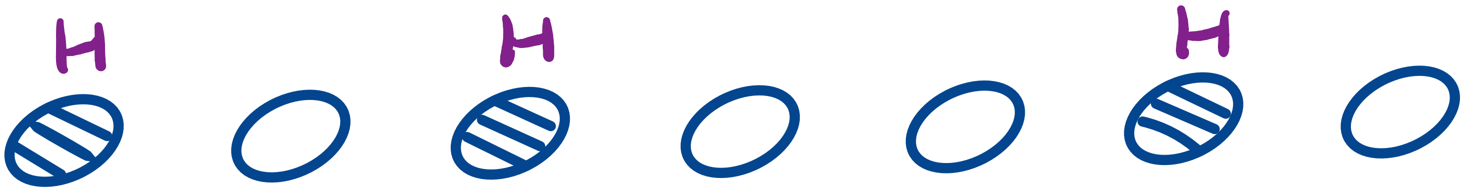
$$X_s = \sum_{i}^{N} X_i \qquad X_i(\omega) = \begin{cases} 1 & \omega \text{ is event } A \\ 0 & \omega \text{ is } A^c \end{cases}$$

Range of X_s is ?
$[0, N]$

Number of heads in N tosses.



Binomial distribution



$N=7$ $\quad P(X_s=3)=?$

Consider random tossing 7 times a biased coin

$$P(X_s=3)=\sum P(\text{any one instance of } 3Hs)$$

$$P(X_s=3) = \binom{7}{3} p^3 (1-p)^4$$

$$\binom{7}{3} P(C_1=H, C_2=T, C_3=H, \ldots)$$

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

$$p^3(1-p)^4$$

Binomial distribution

$N \quad \circ \quad \circ \quad \circ \quad \cdots \quad \circ$

K^{th} heads

$$P(X=K) = \binom{N}{K} p^K (1-p)^{N-K}$$

$$N \geq K \geq 0$$

$$(p+q)^n = \sum \binom{n}{k} p^k q^{n-k} = 1$$

if $q = 1-p$

Binomial distribution

- A discrete random variable X is binomial if

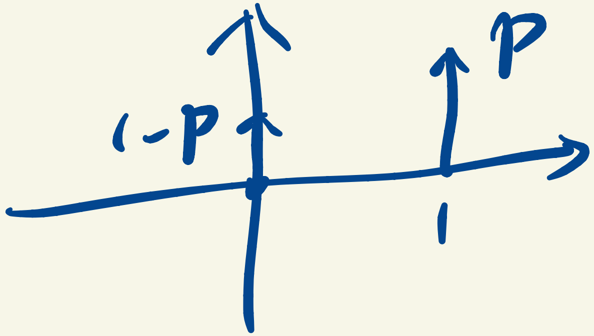
$$P(X = k) = \binom{N}{k} p^k (1-p)^{N-k} \quad for\ integer\ 0 \leq k \leq N$$

with $E[X] = Np \quad \& \quad var[X] = Np(1-p)$

- Examples
  - If we roll a six-sided die N times, how many sixes we will see
  - If I attempt **N** free throws, how many points will I score
  - **What is the sum of N independent and identically distributed Bernoulli trials?**

$$X_s = \sum_{i=1}^{N} X_i$$

$$X_i = \begin{cases} 1 & H \\ 0 & T \end{cases}$$



$$E[X_s] = \sum E[X_i]$$
$$= N \cdot p$$

$$E[X_i] = p$$

$$var[X_s] = var[\sum X_i]$$

$\because$ X_i are indpt

X_i are iid

$$= \sum (var[X_i]) \quad \because iid$$

$$var[X_i] = p(1-p)$$

$$= N \cdot p(1-p)$$

Expectations of Binomial distribution

- A discrete random variable X is binomial if

$$P(X=k) = \binom{N}{k} p^k (1-p)^{N-k} \quad for\ integer\ 0 \le k \le N$$

with $E[X] = Np \quad \& \quad var[X] = Np(1-p)$

Binomial distribution: die example

- Let X be the number of sixes in 36 rolls of a fair six-sided die. What is P(X=k) for k =5, 6, 7

$$P(X=6) = \binom{36}{6}\left(\frac{1}{6}\right)^{6}\left(\frac{5}{6}\right)^{36-6}$$

$p=\frac{1}{6}$

$$P(X\geq 6) = \sum_{k=6}^{36}\binom{36}{k}\left(\frac{1}{6}\right)^{k}\left(\frac{5}{6}\right)^{36-k}$$

$$1-\sum_{k=0}^{5} \cdots$$

- Calculate E[X] and var[X]

$$E[X]=36\times\frac{1}{6}=6 \qquad var[X]=36\times\frac{1}{6}\cdot\frac{5}{6}=5$$

Geometric distribution

- A discrete random variable X is geometric if

$$P(X = k) = (1-p)^{k-1}p \qquad k \geq 1$$

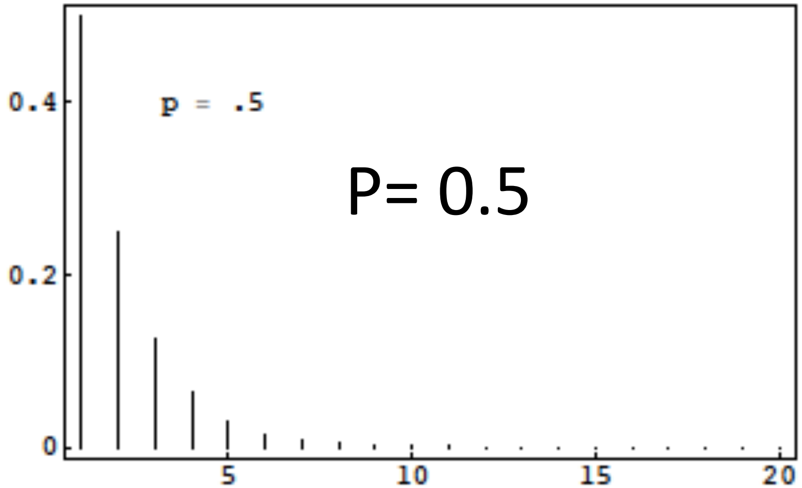
H, TH, TTH, TTTH, TTTTH, $\underbrace{\overbrace{\text{TTTTT}}^{k-1}\text{H}}_{k}$,...

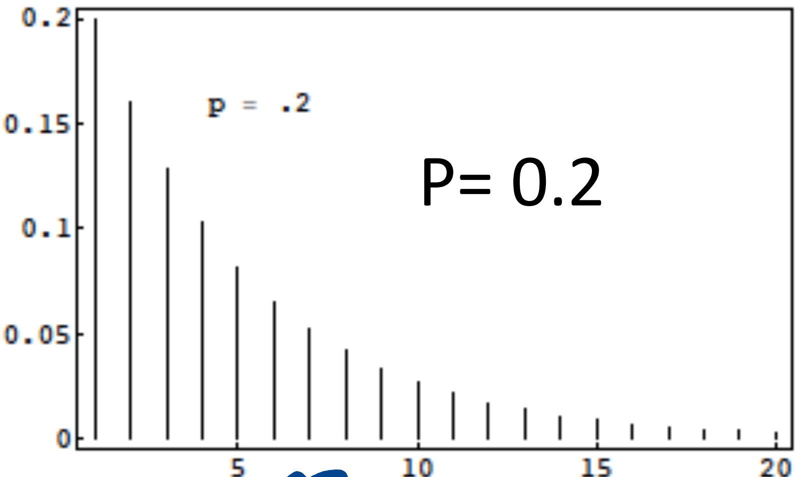
- Expected value and variance

$$E[X] = \frac{1}{p} \quad \& \quad var[X] = \frac{1-p}{p^2}$$

Geometric distribution

$$P(X = k) = (1-p)^{k-1}p \qquad k \geq 1$$





$$E[X] = \sum_{x=1}^{\infty} x\,p(x) = \frac{1}{p}$$

Credit: Prof. Grinstead

Geometric distribution

- Examples:
  - How many rolls of a six-sided die will it take to see the first 6?
  - How many Bernoulli trials must be done before the first 1?
  - How many experiments needed to have the first success?
  - Plays an important role in the **theory of queues**

Derivation of geometric expected value

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

Derivation of geometric expected value

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

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

$\alpha < 1$

$\sum_{n} n\alpha^{n} =$

Derivation of geometric expected value

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

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

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

Derivation of geometric expected value

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

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

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

- For we have

this power series:

Derivation of geometric expected value

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

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

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

- For we have this power series:

$$\sum_{n=1}^{\infty} n x^n = \frac{x}{(1-x)^2}; \quad |x| < 1$$

Derivation of geometric expected value

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

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

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

$$x = 1-p$$

- For we have

this power series:

$$\sum_{n=1}^{\infty} nx^n = \frac{x}{(1-x)^2}; \quad |x| < 1$$

Derivation of geometric expected value

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

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

$$= \frac{p}{1-p}\sum_{k=1}^{\infty} k(1-p)^{k} \quad \boxed{= \frac{1}{p}}$$

- For we have this power series:

$$\sum_{n=1}^{\infty} nx^n = \frac{x}{(1-x)^2}; \qquad |x| < 1$$

Derivation of the power series

$$S(x) = \sum_{n=1}^{\infty} nx^n = \frac{x}{(1-x)^2}; \qquad |x| < 1$$

Proof:

$$\frac{S(x)}{x} = \sum_{n=1}^{\infty} nx^{n-1}\ ; \ \sum_{n=0}^{\infty} x^n = \frac{1}{1-x}; \qquad |x| < 1$$

$$\int_0^x \frac{S(t)}{t} = \sum_{n=1}^{\infty} x^n = x \cdot \frac{1}{1-x} = \frac{x}{1-x}$$

$$\frac{S(x)}{x} = \left(\frac{x}{1-x}\right)'$$

$$S(x) = \frac{x}{(1-x)^2}$$

Geometric distribution: die example

- Let X be the number of rolls of a fair six-sided die needed to see the first 6. What is $P(X = k)$ for k = 1, 2?

$P(X=1) = \frac{1}{6}$

$P(X=2) = \frac{5}{6} \cdot \frac{1}{6}$

T H with $p = \frac{1}{6}$

- Calculate E[X] and var[X]

$$E[X] = \frac{1}{p} \quad \& \quad var[X] = \frac{1-p}{p^2}$$

Betting brainteaser

- What would you rather bet on?
  - How many rolls of a fair six-sided die will it take to see the first 6?
  - How many sixes will appear in 36 rolls of a fair six-sided die?

- Why?

review on your own

Multinomial distribution

- A discrete random variable X is Multinomial if

$$P(X_1 = n_1, X_2 = n_2, ..., X_k = n_k) = \frac{N!}{n_1!n_2!...n_k!} p_1^{n_1} p_2^{n_2} ... p_k^{n_k}$$

$$where \;\; N = n_1 + n_2 + ... + n_k$$

- The event of throwing *N* times the k-sided die to see the probability of getting n_1 X_1, n_2 X_2, n_3 X_3...n_k X_k

A T C G

in DNA

modeling

Multinomial distribution

- A discrete random variable X is Multinomial if

$$P(X_1 = n_1, X_2 = n_2, ..., X_k = n_k) = \frac{N!}{n_1!n_2!...n_k!} p_1^{n_1} p_2^{n_2} ... p_k^{n_k}$$

$$where \;\; N = n_1 + n_2 + ... + n_k$$

- The event of throwing k-sided die to see the probability of getting n_1 X_1, n_2 X_2, n_3 X_3...

ILLINOIS?

$$\frac{8!}{3!2!1!1!1!}$$

I L

Multinomial distribution

- Examples
  - If we roll a six-sided die N times, how many of each value will we see?
  - What are the counts of N independent and identical distributed trials?
  - This is very widely used in genetics

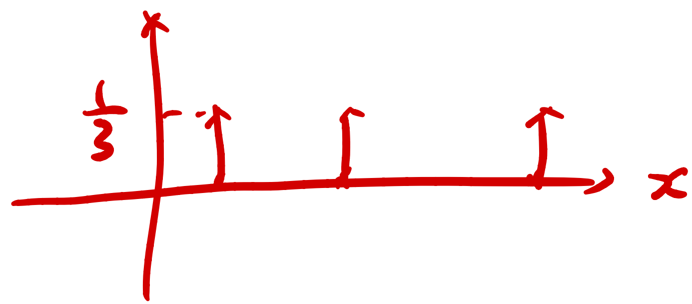
Multinomial distribution: die example

- What is the probability of seeing 1 one, 2 twos, 3 threes, 4 fours, 5 fives and 0 sixes in 15 rolls of a fair six-sided die?

Discrete uniform distribution

- A discrete random variable X is uniform if it takes k different values and

$$P(X = x_i) = \frac{1}{k}$$ For all x_i that X can take



- For example:
  - Rolling a fair k-sided die
  - Tossing a fair coin (k=2)

Discrete uniform distribution

- Expectation of a discrete random variable X that takes k different values uniformly

$$E[X] = \frac{1}{k}\sum_{i=1}^{k} x_i$$

- Variance of a uniformly distributed random variable X .

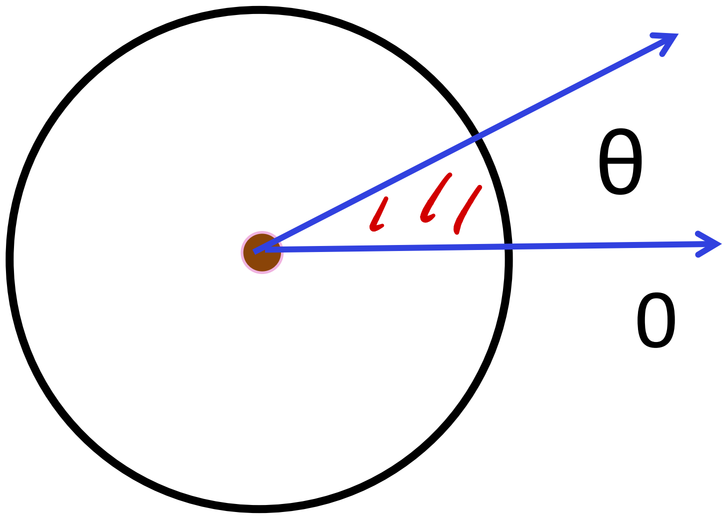
$$var[X] = \frac{1}{k}\sum_{i=1}^{k}(x_i - E[X])^2$$

Example of a continuous random variable

- The spinner

$P(X=k)$

heads



$$\theta \in (0, 2\pi]$$

- The sample space for all outcomes is not countable

$P(\theta = \theta_0) = ?$

$$P(0 < \theta < \theta_0)$$

$$\cancel{P(\theta = \theta_0)}$$

Probability density function (pdf)

- For a continuous random variable X, the probability that X=x is essentially zero for all (or most) x, so we can't define $P(X = x)$
- Instead, we define the **probability density function** (pdf) over an infinitesimally small interval dx, $p(x)dx = P(X \in [x, x + dx])$

$p(x) = \lim_{dx \to 0} \frac{P(x_0 < X < x_0 + dx)}{dx}$

- For $a < b$ $$\int_a^b p(x)dx = P(X \in [a, b])$$

Properties of the probability density function

- $p(x)$ **resembles** the probability function of discrete random variables in that
  - $p(x) \geq 0 \quad$ for all x
  - The probability of X taking all possible values is 1.

$$\int_{-\infty}^{\infty} p(x)dx = 1$$

$= P(X < \infty$
$\& \; X > -\infty) = 1$

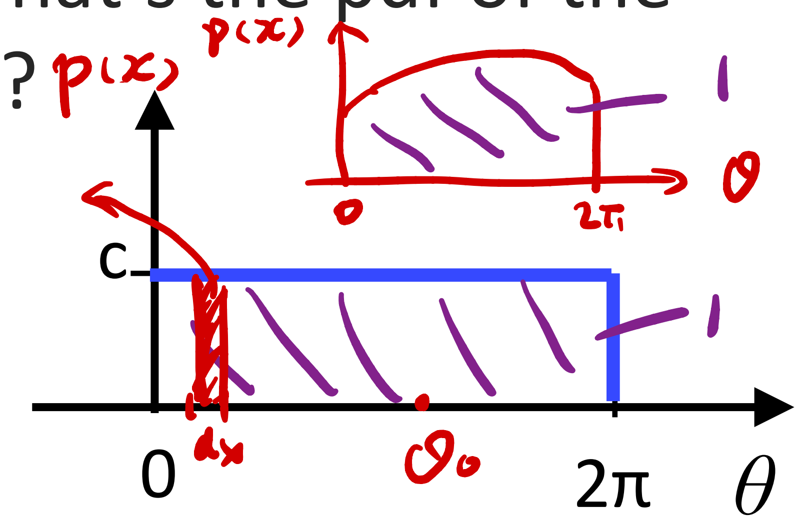
Properties of the probability density function

- $p(x)$ **differs** from the probability distribution function for a discrete random variable in that
  - $p(x)$ is not the probability that $X = x$
  - $p(x)$ can exceed 1

Probability density function: spinner

- Suppose the spinner has equal chance stopping at any position. What's the pdf of the angle θ of the spin position?

$$p(\theta) = \begin{cases} c & if\ \theta \in (0, 2\pi] \\ 0 & otherwise \end{cases}$$



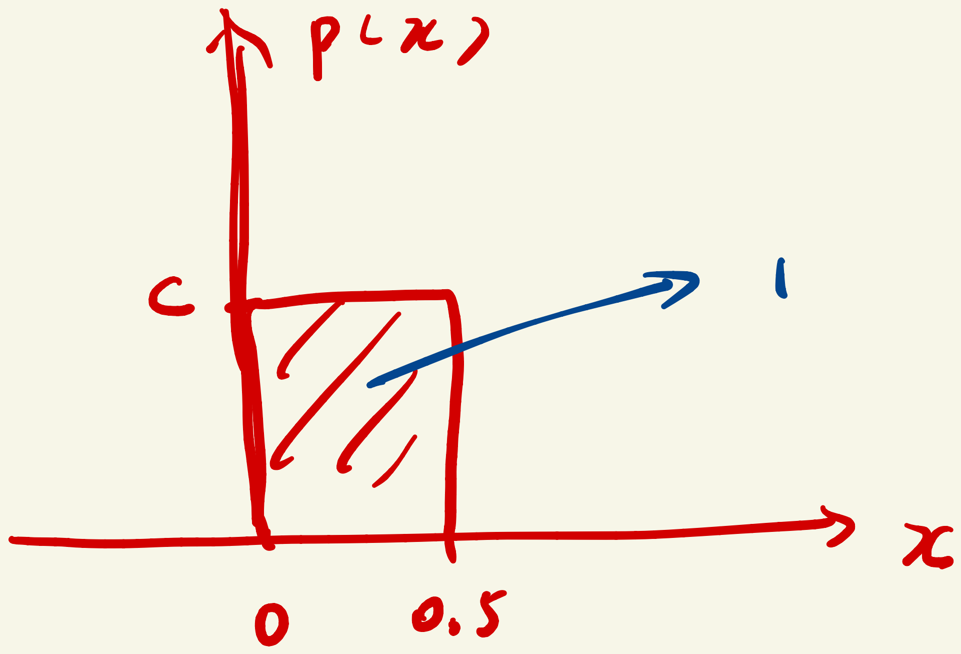
- For this function to be a pdf,

Then $\int_{-\infty}^{\infty} p(\theta) d\theta = 1$

$P(\theta < \theta_0)$

$= \int_0^{\theta_0} p(x)\,dx$

$c = 2$



We ask for $P(X \in [a, b])$

$P(X = x_0) \rightarrow$ not for continuous RV

Assignments

- Work on Week5 material
- Next time: more classic known probability distributions

Additional References

- Charles M. Grinstead and J. Laurie Snell "Introduction to Probability”
- Morris H. Degroot and Mark J. Schervish "Probability and Statistics”

See you next time

See You!