Today

Random Variables.

Review:Poll.

What's an event?

(A) Party at Rao's house.
(B) A protest at Sproul Plaza.
(C) A subset of Ω where Ω is a sample space.
(D) Has a probability associated with it.
(E) Having 2 heads in 3 coin flips.

C,D,E

Bayes Rule is

(A) Awesome.
(B) Allows one to reason from evidence.
(C) $Pr[A|B] = Pr[A \cap B]/Pr[B]$ for events A and B.
(D) Follows from the definition of $Pr[A|B]$.
(E) Converts $P[A|B]$ to $P[B|A]$

A,B,D,E

C is definition of conditional probability

Quick Review: Probability. Some Rules.

- **Sample Space:** Set of outcomes, Ω.
- **Probability:** $Pr[\omega]$ for all $\omega \in \Omega$.
  - $0 \leq Pr[\omega] \leq 1.$, $\sum_{\omega \in \Omega} Pr[\omega] = 1.$
- **Event:** $A \subseteq \Omega$. $Pr[A] = \sum_{\omega \in A} Pr[\omega]$
  - Inclusion/Exclusion: $Pr[A \cup B] = Pr[A] + Pr[B] - Pr[A \cap B]$.
  - Simple Total Probability: $Pr[B] = Pr[A \cap B] + Pr[\overline{A} \cap B]$.
  - Complement: $Pr[\overline{A}] = 1 - Pr[A]$.
  - Union Bound. $Pr[\cup_i A_i] \leq \sum_i Pr[A_i]$
  - Total Probability: $Pr[B] = \sum_i Pr[B \cap A_i]$, for partition $\{A_i\}$.
- **Conditional Probability:** $Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]}$
- Bayes' Rule: $Pr[A|B] = Pr[B|A]Pr[A]/Pr[B]$
  $Pr[A_m|B] = p_m q_m / (\sum_{i=0}^{m} p_i q_i)$, $p_m = Pr[A_m]$, $q_m = Pr[B|A_m]$.
- Product Rule or Intersection Rule:
  $Pr[A_1 \cap \cdots \cap A_n] = Pr[A_1]Pr[A_2|A_1] \cdots Pr[A_n|A_1 \cap \cdots \cap A_{n-1}]$.
- Total Probability/Product: $Pr[B] = Pr[B|A]Pr[A] + Pr[B|\overline{A}]Pr[\overline{A}]$.

Random Variables

Random Variables

1. Random Variables.
2. Expectation
3. Distributions.

Questions about outcomes ...

Experiment: roll two dice.
Sample Space: $\{(1,1),(1,2),\ldots,(6,6)\} = \{1,\ldots,6\}^2$
How many pips? $X((1,1)) = 2$, $X((3,4)) = 7,\ldots$.

Experiment: flip 100 coins.
Sample Space: $\{HHH\cdots H, THH\cdots H,\ldots,TTT\cdots T\}$
How many heads in 100 coin tosses?

Experiment: choose a random student in cs70.
Sample Space: $\{Adam, Jin, Bing,\ldots, Angeline\}$
What midterm score?

Experiment: hand back assignments to 3 students at random.
Sample Space: $\{123, 132, 213, 231, 312, 321\}$
How many students get back their own assignment?

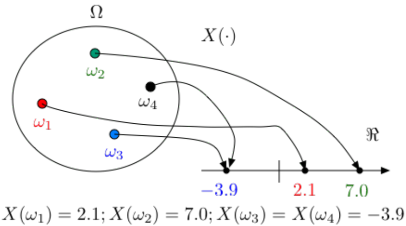
In each scenario, each outcome gives a number.

The number is a (known) function of the outcome.

Random Variables.

A **random variable,** X**,** for an experiment with sample space Ω is a function $X : \Omega \to \Re$.

Thus, $X(\cdot)$ assigns a real number $X(\omega)$ to each $\omega \in \Omega$.



Function $X(\cdot)$ defined on outcomes Ω.

Function $X(\cdot)$ is not random, not a variable!

What varies at random (among experiments)? The outcome!

Random variable makes partition: $A_y = \{\omega \in \Omega : X(\omega) = y\} = X^{-1}(y)$

$A_{-3.9} = \{\omega_3, \omega_4\}, A_{2.1} = \{\omega_1\}, A_{7.0} = \{\omega_2\}$.

Random Variables Definition: Poll.

Random Variable.

(A) Is a function on a sample space.
(B) '$X = i$' is an event in a sample space.
(C) '$X > x$ is an event in a sample space.
(E) For an experiment, $X(\omega)$ is 'random'.
(E) Is neither random nor a variable.

Example 1 of Random Variable

Experiment: roll two dice.
Sample Space: $\{(1,1),(1,2),\ldots,(6,6)\} = \{1,\ldots,6\}^2$
Random Variable X: number of pips.
$X(1,1) = 2$
$X(1,2) = 3$,
$\vdots$
$X(6,6) = 12$,
$X(a,b) = a+b, (a,b) \in \Omega$.

Example 2 of Random Variable

Experiment: flip three coins

Sample Space: $\{HHH, THH, HTH, TTH, HHT, THT, HTT, TTT\}$

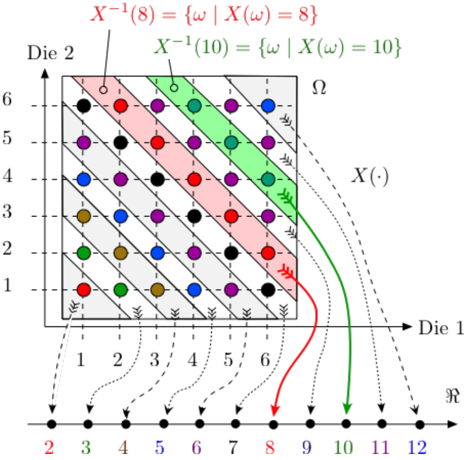
Winnings: if win 1 on heads, lose 1 on tails: X

$X(HHH) = 3$ $X(THH) = 1$ $X(HTH) = 1$ $X(TTH) = -1$

$X(HHT) = 1$ $X(THT) = -1$ $X(HTT) = -1$ $X(TTT) = -3$

Number of pips in two dice.

"What is the likelihood of getting *n* pips?"

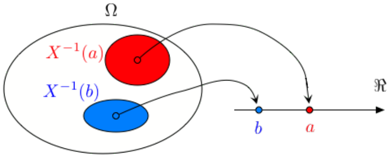


$Pr[X = 10] = 3/36 = Pr[X^{-1}(10)]; Pr[X = 8] = 5/36 = Pr[X^{-1}(8)].$

Distribution

The probability of X taking on a value a.

Definition: The **distribution** of a random variable X, is $\{(a, Pr[X = a]) : a \in \mathscr{A}\}$, where $\mathscr{A}$ is the range of X.



$$Pr[X = a] := Pr[X^{-1}(a)] \text{ where } X^{-1}(a) := \{\omega \mid X(\omega) = a\}.$$

From the "distribution view" is probability space.

(a) $\Omega = \mathscr{A}$

(b) $Pr[a] = Pr[X = a]$

In this part, often connect to experiment with outcomes.

Roll two dice.

How many pips? How many pips on first die? ...

Flip n coins.

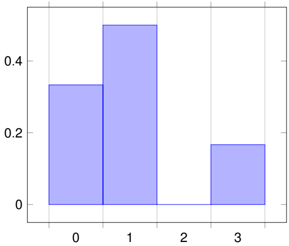
How many heads? How many heads in the first $n/2$ flips? ...

Handing back assignments

Experiment: hand back assignments to 3 students at random.
Sample Space: $\Omega = \{123, 132, 213, 231, 312, 321\}$
How many students get back their own assignment?
Random Variable: values of $X(\omega) : \{3, 1, 1, 0, 0, 1\}$

Distribution:

$$X = \begin{cases} 0, & \text{w.p. } 1/3 \\ 1, & \text{w.p. } 1/2 \\ 3, & \text{w.p. } 1/6 \end{cases}$$



A couple of views.

A probability space: $\Omega = \{0, 1, 3\}$.
$Pr[0] = 1/3, Pr[1] = 1/2, Pr[3] = 1/6$.

"Same" (in a sense) as the distribution of number of fixed points on a permutation of size 3.

Experiment: Can define a random variable (or many) based on a function of any sample space.

Distribution: Can define a sample space using the possible values of a random variable.

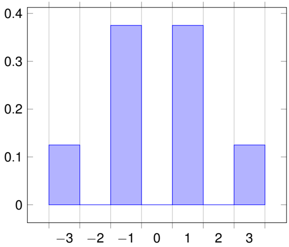
Future:
Continuous distributions: the outcomes are values of random variables.

Flip three coins

Experiment: flip three coins
Sample Space: $\{HHH, THH, HTH, TTH, HHT, THT, HTT, TTT\}$
Winnings: if win 1 on heads, lose 1 on tails. X
Random Variable: $\{3, 1, 1, -1, 1, -1, -1, -3\}$

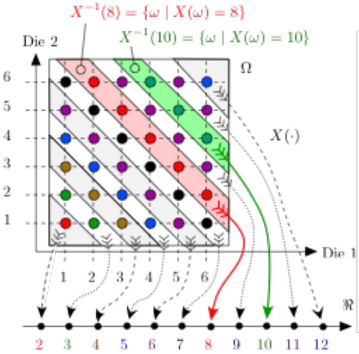
Distribution:

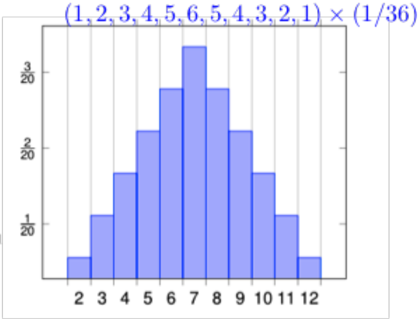
$$X = \begin{cases} -3, & \text{w. p. } 1/8 \\ -1, & \text{w. p. } 3/8 \\ 1, & \text{w. p. } 3/8 \\ 3 & \text{w. p. } 1/8 \end{cases}$$



Number of pips.

Experiment: roll two dice.





Expectation.

How did people do on the midterm?

Distribution.

Summary of distribution?

Average!

Expectation - Intuition

Flip a loaded coin with $Pr[H] = p$ a large number N of times.

Expect heads a fraction p of the times and tails a fraction $1 - p$.

Say that you get 5 for every H and 3 for every T.

With $N(H)$ outcomes with H and $N(T)$ outcomes equal to T, you collect

$$5 \times N(H) + 3 \times N(T).$$

Your average gain per experiment is then

$$\frac{5N(H) + 3N(T)}{N}.$$

Since $\frac{N(H)}{N} \approx p = Pr[X = 5]$ and $\frac{N(T)}{N} \approx 1 - p = Pr[X = 3]$, we find that the average gain per outcome is approximately equal to

$$5Pr[X = 5] + 3Pr[X = 3].$$

We use this frequentist interpretation as a definition.

Expectation - Definition

Definition: The **expected value** of a random variable X is

$$E[X] = \sum_{a \in \mathscr{A}} a \times Pr[X = a].$$

The expected value is also called the mean.

According to our intuition, we expect that if we repeat an experiment a large number N of times and if $X_1, \ldots, X_N$ are the successive values of the random variable, then

$$\frac{X_1 + \cdots + X_N}{N} \approx E[X].$$

That is indeed the case, in the same way that the fraction of times that $X = x$ approaches $Pr[X = x]$.

This (nontrivial) result is called the Law of Large Numbers.

The subjectivist(bayesian) interpretation of $E[X]$ is less obvious.

Expectation: A Useful Fact

Theorem:

$$E[X] = \sum_{\omega} X(\omega) \times Pr[\omega].$$

Proof:

$$\begin{aligned} E[X] &= \sum_{a} a \times Pr[X = a] \\ &= \sum_{a} a \times \sum_{\omega : X(\omega) = a} Pr[\omega] \\ &= \sum_{a} \sum_{\omega : X(\omega) = a} X(\omega) Pr[\omega] \\ &= \sum_{\omega} X(\omega) Pr[\omega] \end{aligned}$$

$\square$

Distributive property of multiplication over addition.

An Example

Flip a fair coin three times.

$\Omega = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$.

$X =$ number of H's: $\{3, 2, 2, 2, 1, 1, 1, 0\}$.

Thus,

$$\sum_{\omega} X(\omega) Pr[\omega] = \{3+2+2+2+1+1+1+0\} \times \frac{1}{8}.$$

Also,

$$\sum_{a} a \times Pr[X = a] = 3 \times \frac{1}{8} + 2 \times \frac{3}{8} + 1 \times \frac{3}{8} + 0 \times \frac{1}{8}.$$

What's the answer? Uh.... $\frac{3}{2}$

Expectation and Average.

There are n students in the class;

$X(m)$ = score of student m, for $m = 1, 2, \ldots, n$.

"Average score" of the n students: add scores and divide by n:

$$\text{Average} = \frac{X(1) + X(1) + \cdots + X(n)}{n}.$$

Experiment: choose a student uniformly at random.
Uniform sample space: $\Omega = \{1, 2, \cdots, n\}, Pr[\omega] = 1/n$, for all ω.
Random Variable: midterm score: $X(\omega)$.
Expectation:

$$E(X) = \sum_{\omega} X(\omega) Pr[\omega] = \sum_{\omega} X(\omega) \frac{1}{n}.$$

Hence,

$$\text{Average } = E(X).$$

This holds for a uniform probability space.

Named Distributions.

Some distributions come up over and over again.

...like “choose” or “stars and bars”....

Let’s cover some.

The binomial distribution: Poll.

Flip n coins with heads probability p.

(A) Number of outcomes is 2^n.
(B) Number of possibilities with k heads $\binom{n}{pk}$.
(C) Probability of k heads is $p^k/2^n$.
(D) Number of possibilities with k heads $\binom{n}{k}$
(E) The probability of every outcome is $1/2^n$.

(A) (D) (E) if $p = 1/2$.

The binomial distribution.

Flip n coins with heads probability p.

Random variable: number of heads.

Binomial Distribution: $Pr[X = i]$, for each i.

How many sample points in event "$X = i$"?
i heads out of n coin flips $\implies \binom{n}{i}$

What is the probability of ω if ω has i heads?
Probability of heads in any position is p.
Probability of tails in any position is $(1-p)$.

$$Pr[\omega] = p^i(1-p)^{n-i}.$$

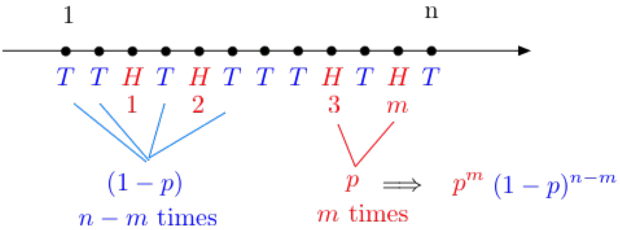
Example: 2 heads/3 flips.
$Pr[HTH] = p \times (1-p) \times p = p^2(1-p)$,
$Pr[THH] = (1-p) \times p \times p = p^2(1-p)$

Probability of "$X = i$" is sum of $Pr[\omega]$, $\omega \in$ "$X = i$".

$Pr[X = i] = \binom{n}{i} p^i(1-p)^{n-i}, i = 0, 1, \ldots, n$: $B(n,p)$ distribution

Example: 2 heads/3 flips. $A = |\{THH, HTH, THH\}| = \binom{3}{2}$
$a \in A, Pr[a] = p^2(1-p)$. $\implies Pr[Heads = 2] = \binom{3}{2} p^2(1-p)$.

The binomial distribution.



$\binom{n}{m}$ outcomes with m Hs and $n - m$ Ts

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

Error channel and...

A packet is corrupted with probability p.

Send $n+2k$ packets.

Probability of at most k corruptions.

$$\sum_{i \leq k} \binom{n+2k}{i} p^i (1-p)^{n+2k-i}.$$

Also distribution in polling, experiments, etc.

Expectation of Binomial Distibution

Parameter p and n. What is the expectation? Guess? pn.

$$Pr[X = i] = \binom{n}{i} p^i (1-p)^{n-i}, i = 0, 1, \ldots, n : B(n, p) \text{ distribution}$$

$$E[X] = \sum_i i \times Pr[X = i].$$

Uh oh? Well... It is pn.

Proof? After linearity of expectation this is easy.

Waiting is good.

Uniform Distribution

Roll a six-sided balanced die. Let X be the number of pips (dots).

X is equally likely to take any of the values $\{1, 2, \ldots, 6\}$.

X is *uniformly distributed* in $\{1, 2, \ldots, 6\}$.

Def: X is *uniformly distributed* in $\{1, 2, \ldots, n\}$ if $Pr[X = m] = 1/n$ for $m = 1, 2, \ldots, n$.

In that case,

$$E[X] = \sum_{m=1}^{n} m Pr[X = m] = \sum_{m=1}^{n} m \times \frac{1}{n} = \frac{1}{n} \frac{n(n+1)}{2} = \frac{n+1}{2}.$$

Geometric Distribution:Poll.

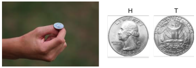
Let's flip a coin with $Pr[H] = p$ until we get H.

The probability of exactly i flips is:

(A) With i flips you have $i-1$ tails and 1 heads.
(B) p^i
(C) $(1-p)^{i-1}p$.
(D) $(1-p)^{i-3}p(1-p)^2$ for $i > 4$.

Geometric Distribution

Let's flip a coin with $Pr[H] = p$ until we get H.



For instance:

$$\begin{aligned}\omega_1 &= H, \text{ or}\\ \omega_2 &= T\ H, \text{ or}\\ \omega_3 &= T\ T\ H, \text{ or}\\ \omega_n &= T\ T\ T\ T\ \cdots\ T\ H.\end{aligned}$$

Note that $\Omega = \{\omega_n, n = 1, 2, \ldots\}$.

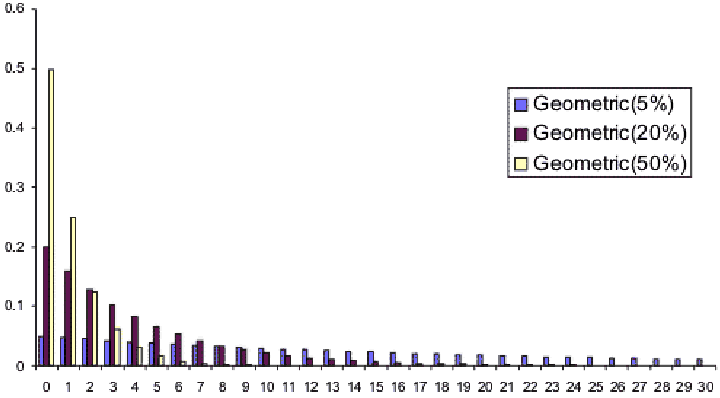
Let X be the number of flips until the first H. Then, $X(\omega_n) = n$.

Also,

$$Pr[X = n] = (1-p)^{n-1}p,\ n \geq 1.$$

Geometric Distribution

$$Pr[X = n] = (1 - p)^{n-1} p, n \geq 1.$$



Oops: $Pr[X - 1]$.

Geometric Distribution

$$Pr[X = n] = (1-p)^{n-1}p, n \geq 1.$$

Note that

$$\sum_{n=1}^{\infty} Pr[X_n] = \sum_{n=1}^{\infty}(1-p)^{n-1}p = p\sum_{n=1}^{\infty}(1-p)^{n-1} = p\sum_{n=0}^{\infty}(1-p)^{n}.$$

Now, if $|a| < 1$, then $S := \sum_{n=0}^{\infty} a^n = \frac{1}{1-a}$. Indeed,

$$\begin{aligned} S &= 1 + a + a^2 + a^3 + \cdots \\ aS &= \quad\; a + a^2 + a^3 + a^4 + \cdots \\ (1-a)S &= 1 + a - a + a^2 - a^2 + \cdots = 1. \end{aligned}$$

Hence,

$$\sum_{n=1}^{\infty} Pr[X_n] = p\,\frac{1}{1-(1-p)} = 1.$$

Geometric Distribution: Expectation

$$X =_D G(p), \text{ i.e., } Pr[X = n] = (1-p)^{n-1}p, n \geq 1.$$

One has

$$E[X] = \sum_{n=1}^{\infty} nPr[X = n] = \sum_{n=1}^{\infty} n \times (1-p)^{n-1}p.$$

Thus,

$$\begin{aligned}
E[X] &= p + 2(1-p)p + 3(1-p)^2p + 4(1-p)^3p + \cdots \\
(1-p)E[X] &= \qquad (1-p)p + 2(1-p)^2p + 3(1-p)^3p + \cdots \\
pE[X] &= p + \ (1-p)p + \ (1-p)^2p + \ (1-p)^3p + \cdots \\
& \qquad \qquad \text{by subtracting the previous two identities} \\
&= \sum_{n=1}^{\infty} Pr[X = n] = 1.
\end{aligned}$$

Hence,

$$E[X] = \frac{1}{p}.$$

Poisson: Motivation and derivation.

McDonalds: How many McDonalds customers arrive in an hour?

Know: average is λ.

What is distribution?

Example: $Pr[2\lambda$ arrivals $]$?

Assumption: "arrivals are independent."

Derivation: cut hour into n intervals of length $1/n$.
$Pr[$ two arrivals $]$ is "$(\lambda/n)^2$" or small if n is large.
Model with binomial.

Poisson

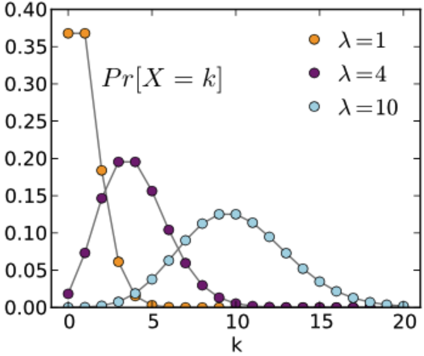
Experiment: flip a coin n times.

The coin is such that $Pr[H] = \lambda/n$.

Random Variable: X - number of heads.

Thus, $X = B(n, \lambda/n)$.

Poisson Distribution is distribution of X "for large n."



Poisson

Experiment: flip a coin n times. The coin is such that $Pr[H] = \lambda/n$.
Random Variable: X - number of heads. Thus, $X = B(n, \lambda/n)$.
Poisson Distribution is distribution of X "for large n."
We expect $X \ll n$. For $m \ll n$ one has

$$
\begin{aligned}
Pr[X = m] &= \binom{n}{m} p^m (1-p)^{n-m}, \text{ with } p = \lambda/n \\
&= \frac{n(n-1)\cdots(n-m+1)}{m!} \left(\frac{\lambda}{n}\right)^m \left(1 - \frac{\lambda}{n}\right)^{n-m} \\
&= \frac{n(n-1)\cdots(n-m+1)}{n^m} \frac{\lambda^m}{m!} \left(1 - \frac{\lambda}{n}\right)^{n-m} \\
&\approx^{(1)} \frac{\lambda^m}{m!} \left(1 - \frac{\lambda}{n}\right)^{n-m} \approx^{(2)} \frac{\lambda^m}{m!} \left(1 - \frac{\lambda}{n}\right)^{n} \approx \frac{\lambda^m}{m!} e^{-\lambda}.
\end{aligned}
$$

For (1) we used $m \ll n$; for (2) we used $(1 - a/n)^n \approx e^{-a}$.

Poisson Distribution: Definition and Mean

Definition Poisson Distribution with parameter $\lambda > 0$

$$X = P(\lambda) \Leftrightarrow Pr[X = m] = \frac{\lambda^m}{m!} e^{-\lambda}, m \geq 0.$$

Fact: $E[X] = \lambda$.

Proof:

$$\begin{aligned} E[X] &= \sum_{m=1}^{\infty} m \times \frac{\lambda^m}{m!} e^{-\lambda} = e^{-\lambda} \sum_{m=1}^{\infty} \frac{\lambda^m}{(m-1)!} \\ &= e^{-\lambda} \sum_{m=0}^{\infty} \frac{\lambda^{m+1}}{m!} = e^{-\lambda} \lambda \sum_{m=0}^{\infty} \frac{\lambda^m}{m!} \\ &= e^{-\lambda} \lambda e^{\lambda} = \lambda. \end{aligned}$$

Second line: Taylor's expansion of e^{λ}. $\square$

Simeon Poisson

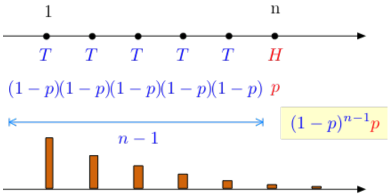
The Poisson distribution is named after:

Siméon Poisson

Siméon Denis Poisson (1781–1840)

Equal Time: B. Geometric

The geometric distribution is named after:



I could not find a picture of D. Binomial, sorry.

Summary

Random Variables

- A random variable X is a function $X : \Omega \to \Re$.
- $Pr[X = a] := Pr[X^{-1}(a)] = Pr[\{\omega \mid X(\omega) = a\}]$.
- $Pr[X \in A] := Pr[X^{-1}(A)]$.
- The distribution of X is the list of possible values and their probability: $\{(a, Pr[X = a]), a \in \mathscr{A}\}$.
- $E[X] := \sum_a a Pr[X = a]$.
- $B(n,p), U[1:n], G(p), P(\lambda)$.