Probability and Statistic (for CS Students)

Lecture 11: Other Discrete Distributions

University of Mazandaran

Fall 1405

Recap: Bernoulli and Binomial Random Variables

The number of successes, in n independent trials, where each trial is a Bernoulli experiment with success probability p:

$$X \sim \text{Bin}(n, p)$$

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Recap: Bernoulli and Binomial Random Variables

Binomial Random Variable

Notation: $X \sim \text{Bin}(n, p)$

Description: Number of "successes" in n identical, independent experiments each with probability of success p.

Parameters: $n \in \{0, 1, \ldots\}$, the number of experiments.
$p \in [0, 1]$, the probability that a single experiment gives a "success".

Support: $x \in \{0, 1, \ldots, n\}$

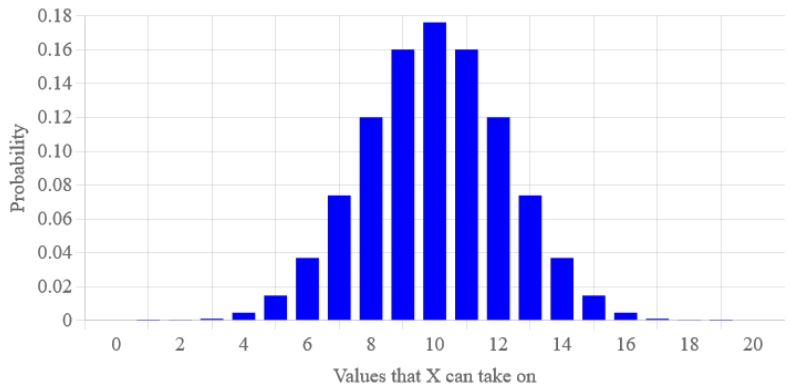
PMF equation: $\mathrm{P}(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$

Expectation: $\mathrm{E}[X] = n \cdot p$

Variance: $\mathrm{Var}(X) = n \cdot p \cdot (1-p)$

PMF graph:

Parameter n: 20 Parameter p: 0.5



Bernoulli Random Variable

Notation: $X \sim \text{Bern}(p)$

Description: A boolean variable that is 1 with probability p

Parameters: p, the probability that $X = 1$.

Support: x is either 0 or 1

PMF equation: $\mathrm{P}(X = x) = \begin{cases} p & \text{if } x = 1 \\ 1-p & \text{if } x = 0 \end{cases}$

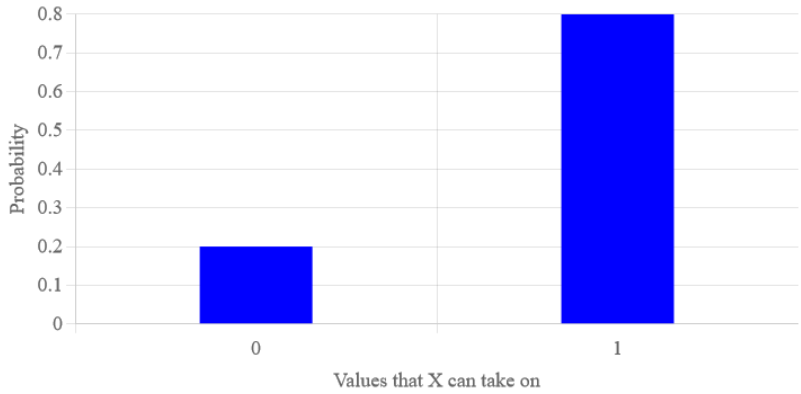
PMF (smooth): $\mathrm{P}(X = x) = p^x (1-p)^{1-x}$

Expectation: $\mathrm{E}[X] = p$

Variance: $\mathrm{Var}(X) = p(1-p)$

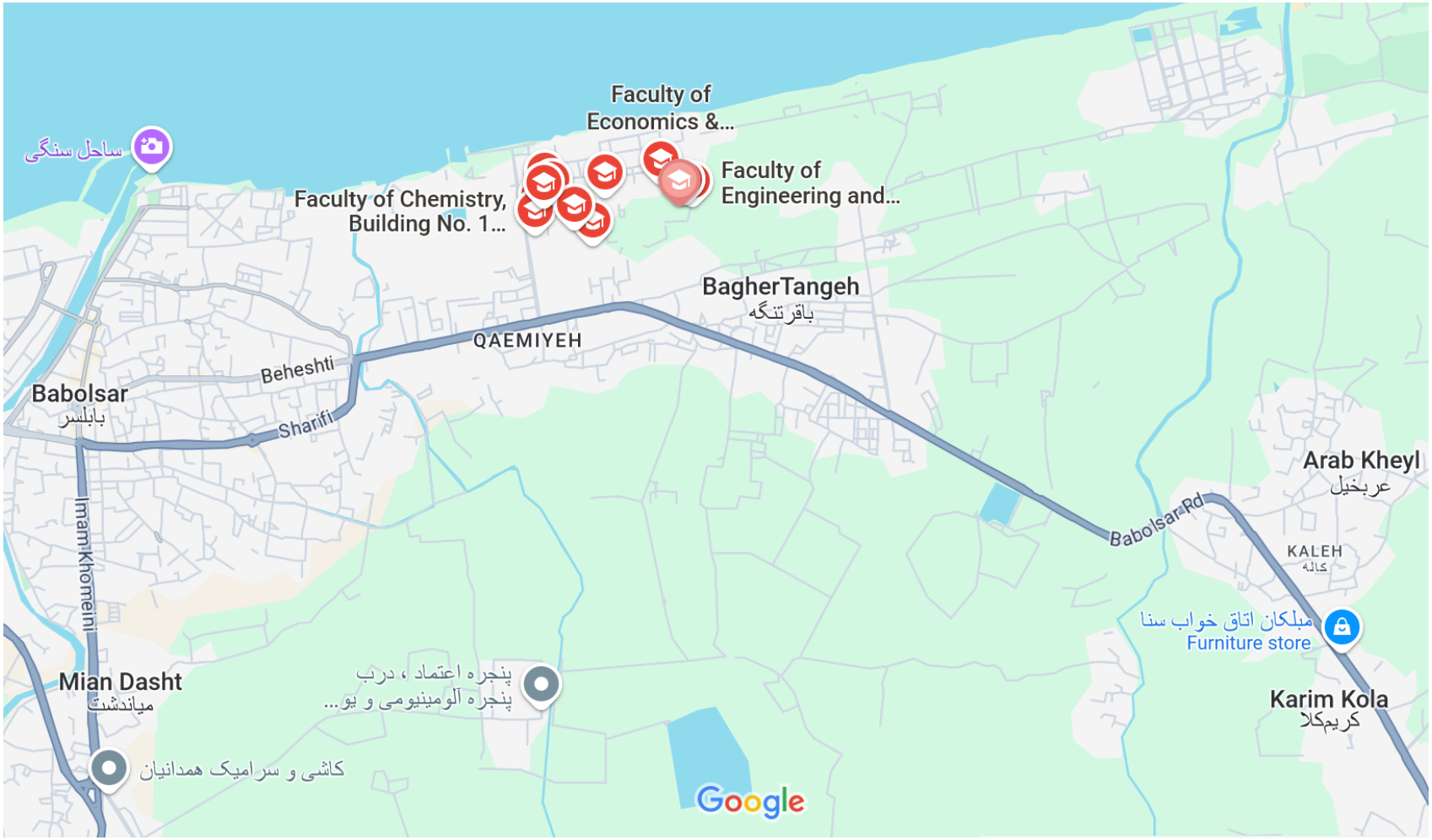
PMF graph:

Parameter p: 0.80



Recap: Poisson Random Variable

Probability of k ride request from this area in the next 1 min



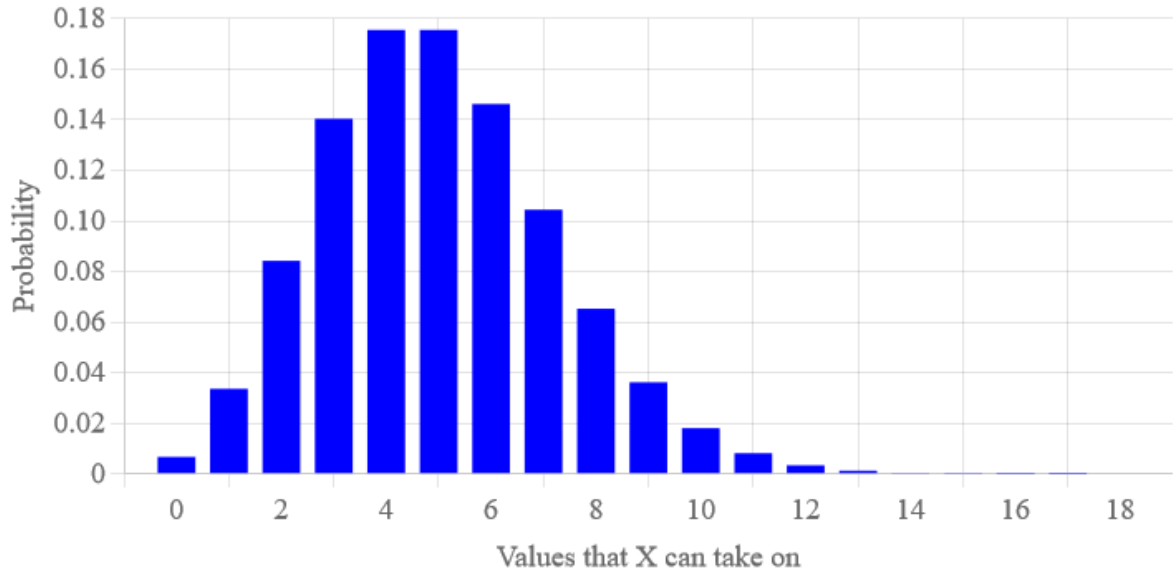
Recap: Poisson Random Variable

Poisson Random Variable

Notation:	$X \sim \text{Poi}(\lambda)$
Description:	Number of events in a fixed time frame if (a) the events occur with a constant mean rate and (b) they occur independently of time since last event.
Parameters:	$\lambda \in \mathbb{R}^+$, the constant average rate.
Support:	$x \in \{0, 1, \ldots\}$
PMF equation:	$\text{P}(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}$
Expectation:	$\text{E}[X] = \lambda$
Variance:	$\text{Var}(X) = \lambda$

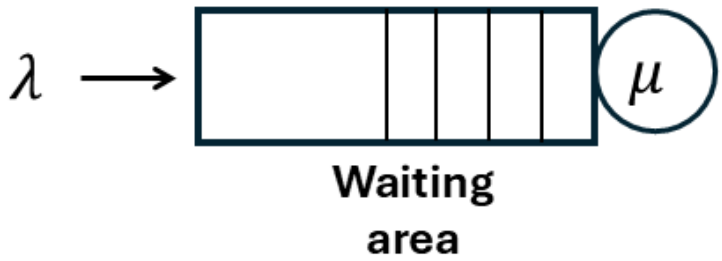
PMF graph:

Parameter λ: 5



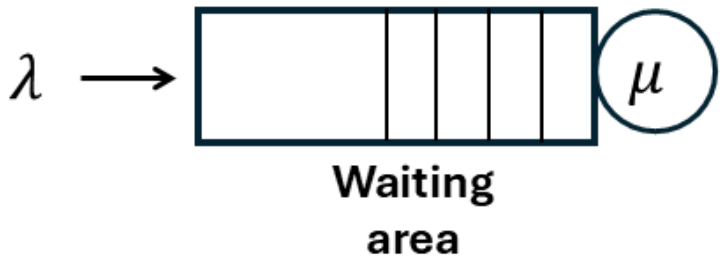
Servers and Their Request Waiting Time

A server wants to know how long a request should wait in the waiting area before it gets served.



Servers and Their Request Waiting Time

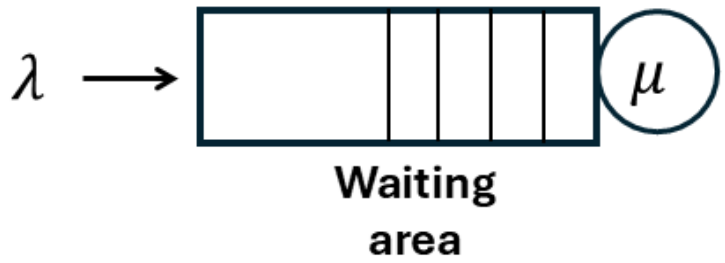
A server wants to know how long a request should wait in the waiting area before it gets served.



One of the most common algorithms is MD1 queue

Servers and Their Request Waiting Time

A server wants to know how long a request should wait in the waiting area before it gets served.

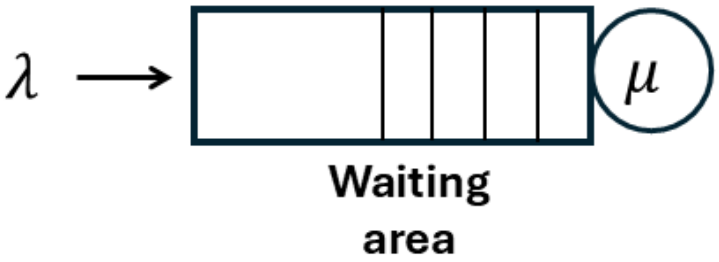


One of the most common algorithms is MD1 queue

You find a paper that says the length of a server "busy period" is distributed as a Borel with parameter $\mu = 0.2$

Servers and Their Request Waiting Time

A server wants to know how long a request should wait in the waiting area before it gets served.



One of the most common algorithms is MD1 queue

You find a paper that says the length of a server "busy period" is distributed as a Borel with parameter $\mu = 0.2$

Borel distribution

3 languages

Article Talk Read Edit View history Tools

From Wikipedia, the free encyclopedia

The **Borel distribution** is a discrete probability distribution, arising in contexts including branching processes and queueing theory. It is named after the French mathematician Émile Borel.

If the number of offspring that an organism has is Poisson-distributed, and if the average number of offspring of each organism is no bigger than 1, then the descendants of each individual will ultimately become extinct. The number of descendants that an individual ultimately has in that situation is a random variable distributed according to a Borel distribution.

Borel distribution

Parameters	$\mu \in [0,1]$
Support	$n \in \{1,2,3,\ldots\}$
PMF	$\frac{e^{-\mu n}(\mu n)^{n-1}}{n!}$
Mean	$\frac{1}{1-\mu}$
Variance	$\frac{\mu}{(1-\mu)^3}$

Definition [edit]

A discrete random variable X is said to have a Borel distribution[1][2] with parameter $\mu \in [0,1]$ if the probability mass function of X is given by

$$P_\mu(n) = \Pr(X = n) = \frac{e^{-\mu n}(\mu n)^{n-1}}{n!}$$

for $n = 1, 2, 3 \ldots$.

Some More Random Variables

	number of successes	time to get successes	
One Trial	$X \sim \text{Ber}(p)$	$X \sim \text{Geo}(p)$	One Success
	$\uparrow n = 1$	$\uparrow r = 1$	
Several Trials	$X \sim \text{Bin}(n, p)$	$X \sim \text{NegBin}(r, p)$	Several Successes
Interval of Time	$X \sim \text{Poi}(\lambda)$	$X \sim \text{Exp}(\lambda)$	One Success

Geometric Random Variable

X is Geometric Random Variable: $X \sim \text{Geo}(p)$

- X is the number of independent trials until first success
- p is probability of success on each trial
- X takes on values $1, 2, 3, \cdots$ with probability:

Geometric Random Variable

X is Geometric Random Variable: $X \sim \text{Geo}(p)$

- X is the number of independent trials until first success
- p is probability of success on each trial
- X takes on values $1, 2, 3, \cdots$ with probability:

$$P(X = n) = (1 - p)^{n-1} p$$

Geometric Random Variable

X is Geometric Random Variable: $X \sim \text{Geo}(p)$

- X is the number of independent trials until first success
- p is probability of success on each trial
- X takes on values $1, 2, 3, \cdots$ with probability:

$$P(X = n) = (1-p)^{n-1} p$$

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

$$\text{Var}(X) = \frac{1-p}{p^2}$$

Negative Binomial Random Variable

X is Negative Binomial Random Variable: $X \sim \text{NegBin}(r, p)$

- X is number of independent trials until r successes
- p is probability of success on each trial
- X takes on values $r, r+1, r+2, \cdots$ with probability

Negative Binomial Random Variable

X is Negative Binomial Random Variable: $X \sim \text{NegBin}(r, p)$

- X is number of independent trials until r successes
- p is probability of success on each trial
- X takes on values $r, r+1, r+2, \cdots$ with probability

$$P(X = n) = \binom{n-1}{r-1} p^r (1-p)^{n-r}, \text{where } n = r, r+1, \cdots$$

Negative Binomial Random Variable

X is Negative Binomial Random Variable: $X \sim \text{NegBin}(r, p)$

- X is number of independent trials until r successes
- p is probability of success on each trial
- X takes on values $r, r+1, r+2, \cdots$ with probability

$$P(X = n) = \binom{n-1}{r-1} p^r (1-p)^{n-r}, \text{where } n = r, r+1, \cdots$$

$$E[X] = \frac{r}{p}$$

$$\text{Var}(X) = r\frac{1-p}{p^2}$$

Geometric and Negative Binomial

Geometric Random Variable

Notation: $X \sim \text{Geo}(p)$

Description: Number of experiments until a success. Assumes independent experiments each with probability of success p.

Parameters: $p \in [0, 1]$, the probability that a single experiment gives a "success".

Support: $x \in \{1, \ldots, \infty\}$

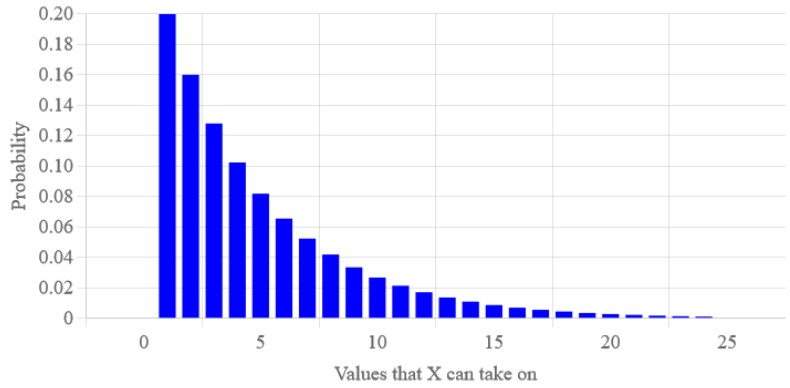
PMF equation: $\text{P}(X = x) = (1-p)^{x-1}p$

Expectation: $\text{E}[X] = \frac{1}{p}$

Variance: $\text{Var}(X) = \frac{1-p}{p^2}$

PMF graph:

Parameter p: 0.20



Negative Binomial Random Variable

Notation: $X \sim \text{NegBin}(r, p)$

Description: Number of experiments until r successes. Assumes each experiment is independent with probability of success p.

Parameters: $r > 0$, the number of success we are waiting for.
$p \in [0, 1]$, the probability that a single experiment gives a "success".

Support: $x \in \{r, \ldots, \infty\}$

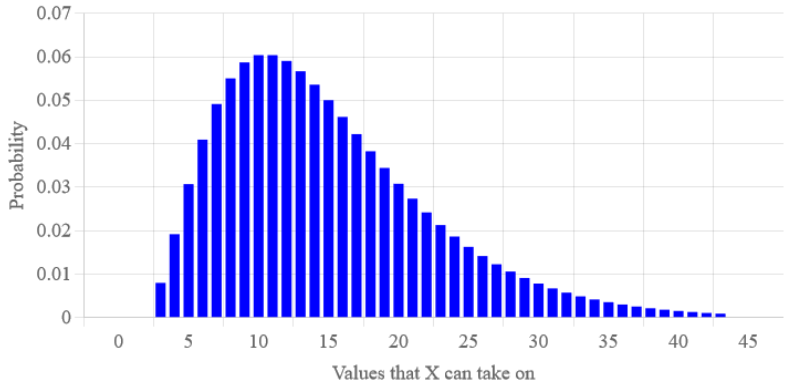
PMF equation: $\text{P}(X = x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}$

Expectation: $\text{E}[X] = \frac{r}{p}$

Variance: $\text{Var}(X) = \frac{r \cdot (1-p)}{p^2}$

PMF graph:

Parameter r: 3 Parameter p: 0.20



Overview of Discrete Distributions

Bernoulli

- indicator of coin flip $X \sim \text{Ber}(p)$

Binomial

- # successes in n coin flips $X \sim \text{Bin}(n, p)$

Poisson

- # successes in a fixed interval of time $X \sim \text{Poi}(\lambda)$

Geometric

- # coin flips until success $X \sim \text{Geo}(p)$

Negative Binomial

- # trials until r successes $X \sim \text{NegBin}(r, p)$

The Challenge of Equity

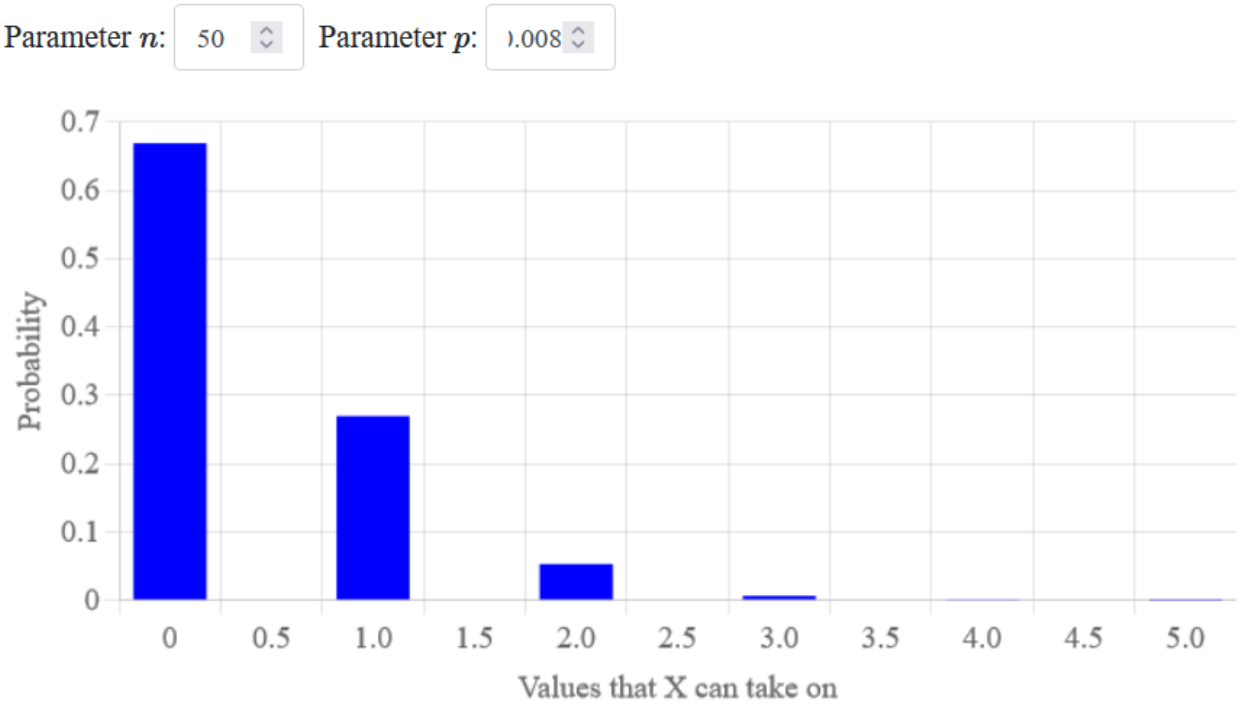
Say, we want to choose a group of 50 students at UMZ and ask them about the most important issues in UMZ to be address asap.

Assume out of 12500 students at UMZ, only 100 are international students. Are we biased against non-Iranian students?

The Challenge of Equity

Say, we want to choose a group of 50 students at UMZ and ask them about the most important issues in UMZ to be address asap.

Assume out of 12500 students at UMZ, only 100 are international students. Are we biased against non-Iranian students?



Bitcoin Mining



SHA-256 Hash(*Data Fixed*, Salt Choice)

Number that looks like random bits

You "mine a bitcoin" if, for given data D, you find a salt number N such that Hash(D, N) produces a string that starts with g zeroes.

Bitcoin Mining

(a) What is the probability that the first number you try will produce a bit string which starts with g zeros (in other words, you mine a bitcoin)

(b) What is the probability that you will need under 100 attempts to mine 2 bit coins?

Bitcoin Mining

(a) What is the probability that the first number you try will produce a bit string which starts with g zeros (in other words, you mine a bitcoin)

Let X be the number of zeros in the first g bits.

(b) What is the probability that you will need under 100 attempts to mine 2 bit coins?

Bitcoin Mining

(a) What is the probability that the first number you try will produce a bit string which starts with g zeros (in other words, you mine a bitcoin)

Let X be the number of zeros in the first g bits.

$X \sim \text{Bin}(n = g, p = 0.5)$

(b) What is the probability that you will need under 100 attempts to mine 2 bit coins?

Bitcoin Mining

(a) What is the probability that the first number you try will produce a bit string which starts with g zeros (in other words, you mine a bitcoin)

Let X be the number of zeros in the first g bits.

$X \sim \text{Bin}(n = g, p = 0.5)$

$P(X = g) = \left(\frac{1}{2}\right)^g$

(b) What is the probability that you will need under 100 attempts to mine 2 bit coins?

Bitcoin Mining

(a) What is the probability that the first number you try will produce a bit string which starts with g zeros (in other words, you mine a bitcoin)

Let X be the number of zeros in the first g bits.

$X \sim \text{Bin}(n = g, p = 0.5)$

$P(X = g) = \left(\frac{1}{2}\right)^g$ call this p_a

(b) What is the probability that you will need under 100 attempts to mine 2 bit coins?

Bitcoin Mining

(a) What is the probability that the first number you try will produce a bit string which starts with g zeros (in other words, you mine a bitcoin)

Let X be the number of zeros in the first g bits.

$X \sim \text{Bin}(n = g, p = 0.5)$

$P(X = g) = \left(\frac{1}{2}\right)^g$ call this p_a

(b) What is the probability that you will need under 100 attempts to mine 2 bit coins?

Let X be

- # bitcoins mined in the first 99 attempts

  $X \sim \text{Bin}(n = 99, p = p_a)$

Bitcoin Mining

(a) What is the probability that the first number you try will produce a bit string which starts with g zeros (in other words, you mine a bitcoin)

Let X be the number of zeros in the first g bits.

$X \sim \text{Bin}(n = g, p = 0.5)$

$P(X = g) = \left(\frac{1}{2}\right)^g$ call this p_a

(b) What is the probability that you will need under 100 attempts to mine 2 bit coins?

Let X be

- # bitcoins mined in the first 99 attempts
  $X \sim \text{Bin}(n = 99, p = p_a)$
- # attempts until 2 bitcoins mined
  $X \sim \text{NegBin}(r = 2, p = p_a)$

Bitcoin Mining

(a) What is the probability that the first number you try will produce a bit string which starts with g zeros (in other words, you mine a bitcoin)

Let X be the number of zeros in the first g bits.

$X \sim \text{Bin}(n = g, p = 0.5)$

$P(X = g) = \left(\frac{1}{2}\right)^g$ call this p_a

(b) What is the probability that you will need under 100 attempts to mine 2 bit coins?

Let X be

- # bitcoins mined in the first 99 attempts
  $X \sim \text{Bin}(n = 99, p = p_a)$
- # attempts until 2 bitcoins mined
  $X \sim \text{NegBin}(r = 2, p = p_a)$

$$P(X < 100) = \sum_{x=2}^{99} P(X = x)$$

Multinomial Random Variable

Binomial RV

What is the probability of getting k successes and $n-k$ failures in n trials?

$$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$$

Multinomial Random Variable

Binomial RV

What is the probability of getting k successes and $n-k$ failures in n trials?

$$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$$

Mutlinomial RV

What is the probability of getting c_1 of outcome 1, c_2 of outcome 2, ..., and c_m of outcome m in n trials?

Multinomial Random Variable

Consider an experiments of n independent trails:

- Each trial results in one of m outcomes. $P\left(\text{outcome } i\right) = p_i$
- Let $X_i = \#$ trials with outcome i

$P\left(X_1 = c_1, X_2 = c_2, \cdots, X_m = c_m\right) =$ $\qquad p_1^{c_1} p_2^{c_2} \cdots p_m^{c_m}$

where $\sum_{i=1}^{m} c_i = n$ and $\sum_{i=1}^{m} p_i = 1$

Multinomial Random Variable

Consider an experiments of n independent trails:

- Each trial results in one of m outcomes. $P\left(\text{outcome } i\right) = p_i$
- Let $X_i = \#$ trials with outcome i

$$P\left(X_1 = c_1, X_2 = c_2, \cdots, X_m = c_m\right) = \binom{n}{c_1, c_2, \cdots, c_m} p_1^{c_1} p_2^{c_2} \cdots p_m^{c_m}$$

where $\sum_{i=1}^{m} c_i = n$ and $\sum_{i=1}^{m} p_i = 1$

Multinomial Random Variable

When there are n objects such that

- n_1 are the same (indistinguishable or indistinct), and
- n_2 are the same, and

  ...
- n_r are the same,

the number of unique orderings (permutations) is

$$\frac{n!}{n_1!n_2!\cdots n_r!} = \binom{n}{n_1, n_2, \cdots, n_r}$$

Multinomial Random Variable

A 6-sided die is rolled 7 times.
What is the probability of getting:

1 one, 0 threes, 0 fives, 1 two, 2 fours, 3 sixes

Multinomial Random Variable

A 6-sided die is rolled 7 times.
What is the probability of getting:

1 one, 0 threes, 0 fives, 1 two, 2 fours, 3 sixes

$X_i = \#$ of trials with outcome i

$P(X_1 = 1, X_2 = 1, X_3 = 0, X_4 = 2, X_5 = 0, X_6 = 3)$

Multinomial Random Variable

A 6-sided die is rolled 7 times.
What is the probability of getting:

1 one, 0 threes, 0 fives, 1 two, 2 fours, 3 sixes

$X_i = \#$ of trials with outcome i

$P(X_1 = 1, X_2 = 1, X_3 = 0, X_4 = 2, X_5 = 0, X_6 = 3)$

$= \binom{7}{1,1,0,2,0,3} \left(\frac{1}{6}\right)^1 \left(\frac{1}{6}\right)^1 \left(\frac{1}{6}\right)^0 \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right)^0 \left(\frac{1}{6}\right)^3 = 420 \left(\frac{1}{6}\right)^7$

Intro to Natural Language Processing

Ignoring the order of words ...

What is the probability of any given word that you write in English?

- $P\left(\text{word} = \text{“the”}\right) > P\left(\text{word} = \text{“pokeman”}\right)$
- $P\left(\text{word} = \text{“Mazandaran”}\right) < P\left(\text{word} = \text{“California”}\right)$

Intro to Natural Language Processing

Ignoring the order of words ...

What is the probability of any given word that you write in English?

- $P\left(\text{word} = \text{“the”}\right) > P\left(\text{word} = \text{“pokeman”}\right)$
- $P\left(\text{word} = \text{“Mazandaran”}\right) < P\left(\text{word} = \text{“California”}\right)$

Probabilities of counts of words = Multinomial distribution

Ignoring the order of words ...

What is the probability of any given word that you write in English?

- $P(\text{word} = \text{“the”}) > P(\text{word} = \text{“pokeman”})$
- $P(\text{word} = \text{“Mazandaran”}) < P(\text{word} = \text{“California”})$

Probabilities of counts of words = Multinomial distribution

A document is a large multinomial (according to the Global Language Monitor, there are 988,968 words in the English language used on the internet!)

Model Test as a Multinomial

Example document:

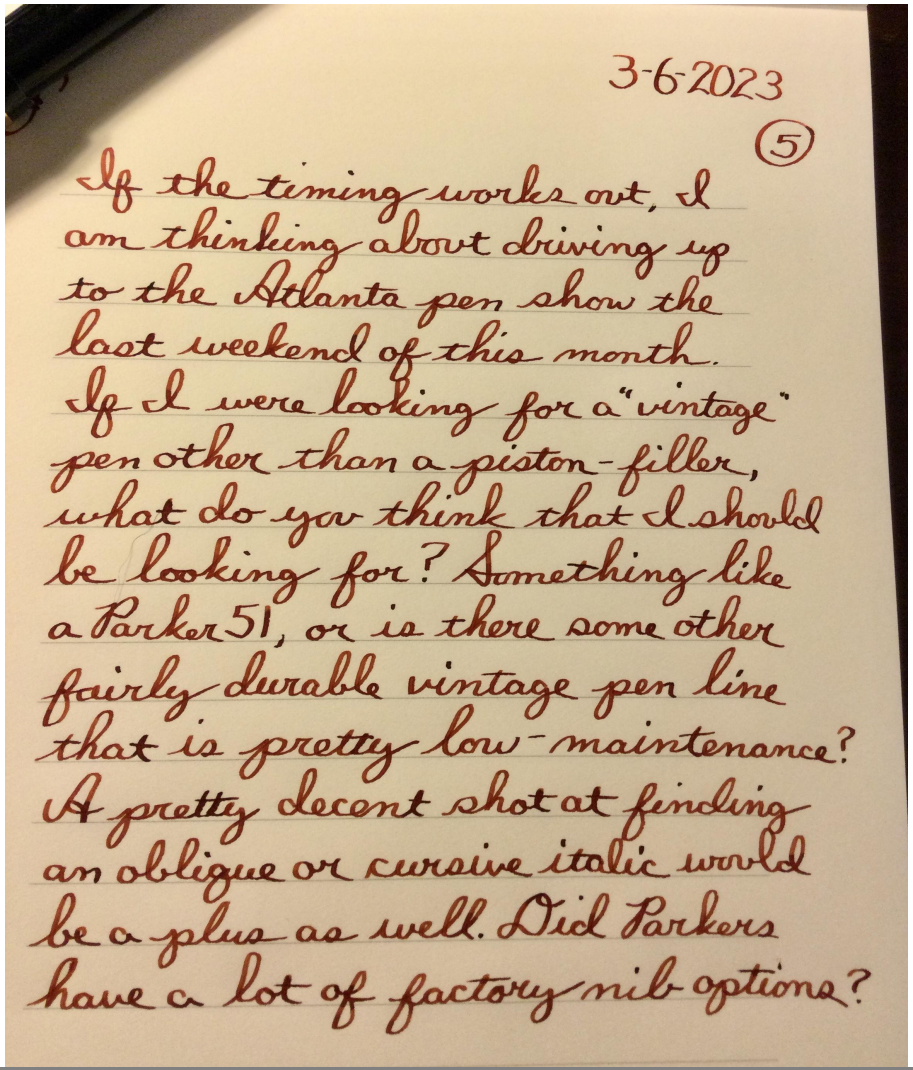
"Pay for butter chicken with a credit-card. Butter chicken is great. So are credit-cards. click for free."

$n = 17$

$$P\left(\text{butter} = 2, \cdots, \text{free} = 1 \middle| \text{spam}\right)$$
$$= \frac{n!}{2!2! \cdots 1!} p_{\text{butter}}^2 p_{\text{for}}^2 p_{\text{chicken}}^2 \cdots p_{\text{free}}^1$$

Who Wrote Which Document?

We had a writing test in our English class. The following writing appears without the writer's name!

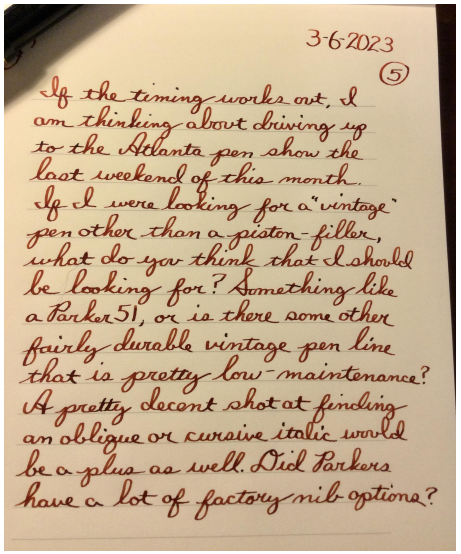


Who Wrote Which Document?

We had a writing test in our English class. The following writing appears without the writer's name!

Two students claim that they wrote the text!

Your task: use your probability knowledge and tell me who is lying!?



Who Wrote Which Document?

We have words, we want to know probability of authorship. We also know probability of words given authors ...

Who Wrote Which Document?

We have words, we want to know probability of authorship. We also know probability of words given authors ...
What we have:

- $P(D|S_i)$: probability of a particular document given each student
- $P(S_i)$: probability of each student write any document

Who Wrote Which Document?

We have words, we want to know probability of authorship. We also know probability of words given authors ...

Who Wrote Which Document?

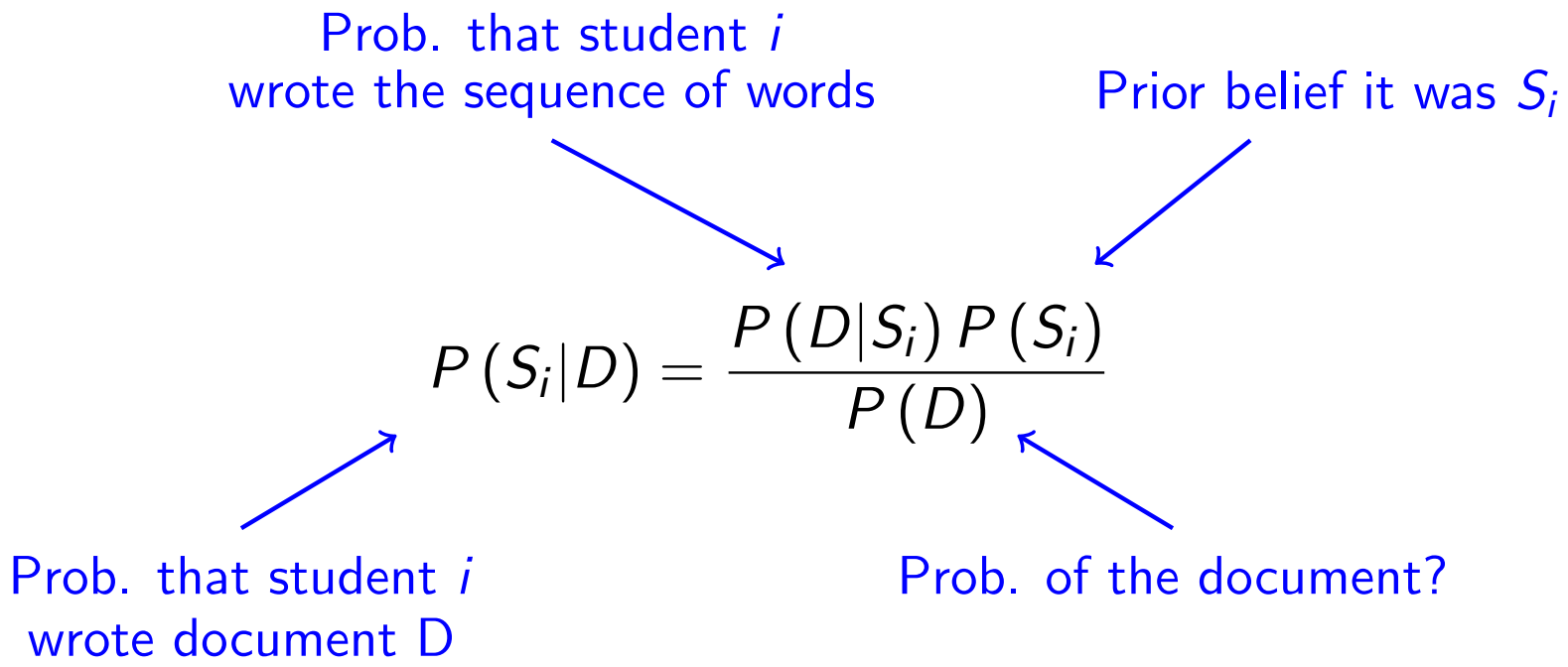
We have words, we want to know probability of authorship. We also know probability of words given authors ...
What we have:

- $P(D|S_i)$: probability of a particular document given each student
- $P(S_i)$: probability of each student write any document

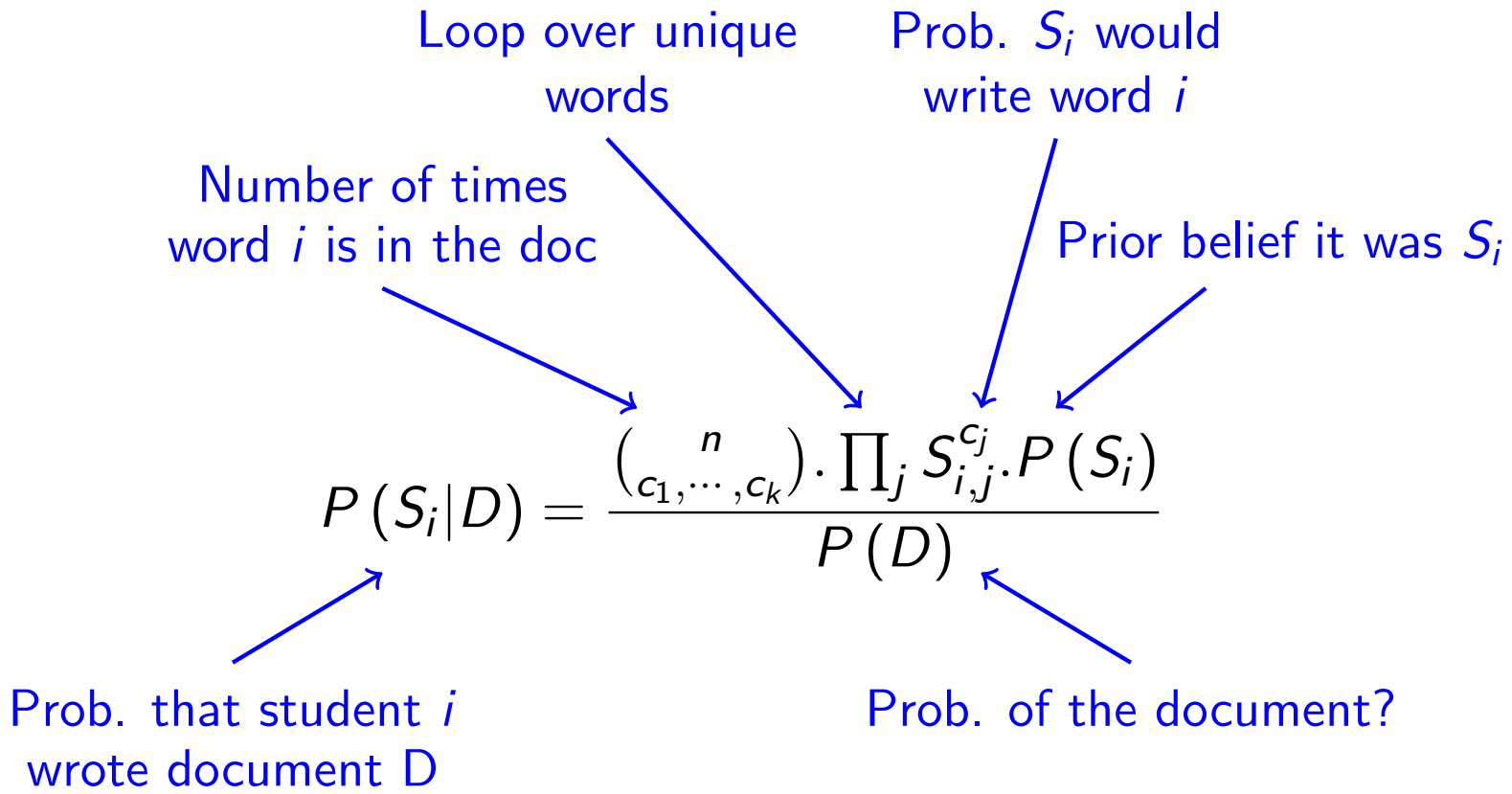
What we want:

- $P(S_i|D)$: probability of a particular student write the document

Who Wrote Which Document?



Who Wrote Which Document?



Who Wrote Which Document?

Prob. that S_1 wrote D Prob. that S_2 wrote D

$$P(S_1|D) = \frac{P(S_1)\binom{n}{c_1,\cdots,c_k}\cdot\prod_j S_{1,j}^{c_j}}{P(D)} \quad P(S_2|D) = \frac{P(S_2)\binom{n}{c_1,\cdots,c_k}\cdot\prod_j S_{2,j}^{c_j}}{P(D)}$$

$$\frac{P(S_1|D)}{P(S_2|D)} = \frac{\frac{P(S_1)\binom{n}{c_1,\cdots,c_k}\cdot\prod_j S_{1,j}^{c_j}}{P(D)}}{\frac{P(S_2)\binom{n}{c_1,\cdots,c_k}\cdot\prod_j S_{2,j}^{c_j}}{P(D)}} = \frac{\prod_j S_{1,j}^{c_j}}{\prod_j S_{2,j}^{c_j}}$$