THE UNIVERSITY OF TEXAS AT AUSTIN
Department of Statistics and Data Sciences
College of Natural Sciences

SDS 321: Introduction to Probability and Statistics

Lecture 8: End of counting and Discrete random variables

Purnamrita Sarkar
Department of Statistics and Data Science
The University of Texas at Austin

www.cs.cmu.edu/∼psarkar/teaching

Probability and counting: example 1b

- A bus with 5 passengers makes 10 stops. All configurations of discharging the passengers are equally likely.
- What is the probability p that no two passengers get down at the same stop?

Probability and counting: example 1b

- A bus with 5 passengers makes 10 stops. All configurations of discharging the passengers are equally likely.
- What is the probability p that no two passengers get down at the same stop?
- # all possible configurations is 10^5.

Probability and counting: example 1b

- A bus with 5 passengers makes 10 stops. All configurations of discharging the passengers are equally likely.
- What is the probability p that no two passengers get down at the same stop?
- \# all possible configurations is 10^5.
- \# configurations with 5 passengers each with a different stop is $(10)_5$.

Probability and counting: example 1b

- A bus with 5 passengers makes 10 stops. All configurations of discharging the passengers are equally likely.
- What is the probability p that no two passengers get down at the same stop?
- # all possible configurations is 10^5.
- # configurations with 5 passengers each with a different stop is $(10)_5$.
- $p = (10)_5 / 10^5$.

Probability and counting: example 1c

- The birthdays of $r \leq 365$ people form a sample of size r from the population of all birthdays (365 days in the year). We assume that a person is equally likely to be born on any of the 365 days and no one was born on Feb 29^{th}.
- What is the probability that no two people will have the same birthday?

Probability and counting: example 1c

- The birthdays of $r \leq 365$ people form a sample of size r from the population of all birthdays (365 days in the year). We assume that a person is equally likely to be born on any of the 365 days and no one was born on Feb 29^{th}.
- What is the probability that no two people will have the same birthday?
- $$p = \frac{(365)_r}{365^r} = \frac{365!}{(365-r)!365^r} \tag{1}$$

Probability and counting: example 1c

- The birthdays of $r \leq 365$ people form a sample of size r from the population of all birthdays (365 days in the year). We assume that a person is equally likely to be born on any of the 365 days and no one was born on Feb 29^{th}.
- What is the probability that no two people will have the same birthday?
- $$p = \frac{(365)_r}{365^r} = \frac{365!}{(365-r)!365^r} \tag{1}$$
- What is the probability if $r = 366$?

Probability and counting: example 1c

- The birthdays of $r \leq 365$ people form a sample of size r from the population of all birthdays (365 days in the year). We assume that a person is equally likely to be born on any of the 365 days and no one was born on Feb 29^{th}.
- What is the probability that no two people will have the same birthday?
- $$p = \frac{(365)_r}{365^r} = \frac{365!}{(365-r)!365^r} \qquad (1)$$
- What is the probability if $r = 366$?
- Then at least two people must have the same birthday. This is also called the Pigeonhole Principle. So $p = 0$.

Probability and counting: example 1c

- The birthdays of $r \leq 365$ people form a sample of size r from the population of all birthdays (365 days in the year). We assume that a person is equally likely to be born on any of the 365 days and no one was born on Feb 29^{th}.
- What is the probability that no two people will have the same birthday?
-

$$p = \frac{(365)_r}{365^r} = \frac{365!}{(365-r)!365^r} \tag{1}$$

- What is the probability if $r = 366$?
- Then at least two people must have the same birthday. This is also called the Pigeonhole Principle. So $p = 0$.
- How will you calculate this for large r, say $r = 30$?

Birthdays

- Problem... my calculator can't handle $365!$ or 365^{30}.
- Take logarithms! $365^{30} = 10^{76.8688} = 7.392 \times 10^{76}$.
- We can *approximate* factorials using Stirling's approximation:

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n,$$

- The $\sim$ symbol means the ratio of the two sides tend to 1 as $n \to \infty$.

$$\begin{aligned}
\ln(365!) &\approx 1792.3 \\
\ln(335!) &\approx 1616.6 \\
\ln\left(\frac{365!}{335!}\right) &= \ln(365!) - \ln(335!) \approx 1792.3 - 1616.6 = 175.55 \\
\frac{365!}{335!} &\approx e^{175.55} = 2.1711 \times 10^{76}
\end{aligned}$$

- The actual value is 2.1710×10^{76} – not bad!

Birthdays

- So, we with 30 people, we have 7.392×10^{76} possible combinations of birthdays.
- 2.171×10^{76} of these possible combinations of birthdays have no repeats.
- So, the probability of no one having the same birthday is:

$$\frac{2.171 \times 10^{76}}{7.392 \times 10^{76}} \approx 0.296$$

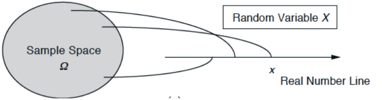
- Odds are, there's a shared birthday!

Random Variables

- So far we have talked about events and sample spaces.
- However for many experiments it is easier to use a summary variable.
- Say we are taking a opinion poll among 100 students about how understandable the lectures are.
- If "1" is used for understandable and "0" is used for not, then there are 2^{100} possible outcomes!!
- However, the thing that matters most is the number of students who think the class is understandable (or equivalently not).
- If we define a variable X to be that number, then the range of X is $\{0, 1, \ldots, 100\}$. **Much easier to handle that!**

Random Variable as a Mapping

Random Variable. A random variable is function from the sample space Ω into the real numbers.



Examples

- You toss a coin: is it head or tail?
- You roll a die: what number do you get?
- Number of heads in three coin tosses
- The sum of two rolls of die
- The number of die rolls it takes to get a six.

Discrete and continuous random variables

- A random variable is discrete if its range (the values it can take) is finite or at most countably finite.
  - $X =$ sum of two rolls of a die. $X \in \{2, \ldots, 12\}$.
  - $X =$ number of heads in 100 coin tosses. $X \in \{0, \ldots, 100\}$
  - $X =$ number of coin tosses to get a head. $X \in \{1, 2, 3, \ldots\}$
- Consider an experiment where you throw a dart which can fall anywhere between $[-1, 1]$.
- Let X be a the dart's position d. X is a random variable and X is not discrete.
- Now let Y be a random variable such that:

$$Y = \begin{cases} 1 & \text{If } d < 0 \\ 0 & \text{If } d \geq 0 \end{cases} \tag{2}$$

Now Y is a discrete random variable.

Probability Mass Functions

Earlier we talked about probability laws which assign probabilities to events in a sample space. One can play the same game with random variables.

- $X =$ number of heads in two fair coin tosses
- X can take values in $\{0, 1, 2\}$.
- $P(X = 0) =$
- $P(X = 1) =$
- $P(X = 2) =$

Probability Mass Functions

Earlier we talked about probability laws which assign probabilities to events in a sample space. One can play the same game with random variables.

- $X =$ number of heads in two fair coin tosses
- X can take values in $\{0, 1, 2\}$.
- $P(X = 0) = P(\{TT\}) = 1/4$.
- $P(X = 1) =$
- $P(X = 2) =$

Probability Mass Functions

Earlier we talked about probability laws which assign probabilities to events in a sample space. One can play the same game with random variables.

- $X =$ number of heads in two fair coin tosses
- X can take values in $\{0, 1, 2\}$.
- $P(X = 0) = P(\{TT\}) = 1/4$.
- $P(X = 1) = P(\{HT, TH\}) = 1/2$.
- $P(X = 2) =$

Probability Mass Functions

Earlier we talked about probability laws which assign probabilities to events in a sample space. One can play the same game with random variables.

- $X =$ number of heads in two fair coin tosses
- X can take values in $\{0, 1, 2\}$.
- $P(X = 0) = P(\{TT\}) = 1/4$.
- $P(X = 1) = P(\{HT, TH\}) = 1/2$.
- $P(X = 2) = P(\{HH\}) = 1/4$.

Probability Mass Functions

Earlier we talked about probability laws which assign probabilities to events in a sample space. One can play the same game with random variables.

- $X =$ number of heads in two fair coin tosses
- X can take values in $\{0, 1, 2\}$.
- $P(X = 0) = P(\{TT\}) = 1/4$.
- $P(X = 1) = P(\{HT, TH\}) = 1/2$.
- $P(X = 2) = P(\{HH\}) = 1/4$.
- In general, the probability that a random variable X takes up a value x is written as $p_X(x)$, or $P_X(x)$ or $P_X(X = x)$ etc.
- A random variable is always written with upper-case and the numerical value we are trying to evaluate the probability for is written with a lower case.

Properties of P.M.F's

- Perhaps not so surprisingly, $\sum_{x} P(X = x) = 1$
- $P(X \in S) = \sum_{x \in S} P(X = x)$
- To compute $P(X = x)$
  - Collect all possible outcomes that give $\{X = x\}$.
  - Add their probabilities to get $P(X = x)$.
- You are throwing 2 fair coins. What is the probability that you see at least one head?
  - Well at least one head is $\{HT, TH, HH\}$. So probability of this is $3/4$.

The Uniform random variable

Consider the roll of a fair die. You are interested in the number on the roll.

- X can take how many different values?
- Now what are the probabilities of taking on those values?
- $P(X = 1) = P(X = 2) = P(X = 3) = P(X = 4) = P(X = 5) = P(X = 6) = 1/6$
- For an **uniform** random variable X, each value has equal probability mass.
- If an **uniform** random variable X takes on k different values, then the probability mass at each of those values are $1/k$.

The Bernoulli random variable

Consider the toss of a biased coin, which gives a head with probability p.

- A **Bernoulli** random variable X takes two values: 1 if a head comes up and 0 if not. $X = \begin{cases} 1 & \text{If head} \\ 0 & \text{If tail} \end{cases}$
- The PMF is given by: $p_X(x) = \begin{cases} p & \text{If } x = 1 \\ 1 - p & \text{If } x = 0 \end{cases}$
- Examples of a Bernoulli
  - A person can be healthy or sick with a certain disease.
  - A test result for a disease can be positive or negative.
  - It may rain one day or not.

Example I

You are throwing 10 fair coins. What is the probability that the sum X equals 5?

- Count all binary strings of length 10 with 5 1's.
- We did this last time!

Example I

You are throwing 10 fair coins. What is the probability that the sum X equals 5?

- Count all binary strings of length 10 with 5 1's.
- We did this last time! $\binom{10}{5}$.

Example I

You are throwing 10 fair coins. What is the probability that the sum X equals 5?

- Count all binary strings of length 10 with 5 1's.
- We did this last time! $\binom{10}{5}$.
- Probability of any one of these events is $(1/2)^{10}$.

Example I

You are throwing 10 fair coins. What is the probability that the sum X equals 5?

- Count all binary strings of length 10 with 5 1's.
- We did this last time! $\binom{10}{5}$.
- Probability of any one of these events is $(1/2)^{10}$.
- So $P(X = 5) = \frac{\binom{10}{5}}{2^{10}}$

Example I

You are throwing 10 fair coins. What is the probability that the sum X equals 5?

- Count all binary strings of length 10 with 5 1's.
- We did this last time! $\binom{10}{5}$.
- Probability of any one of these events is $(1/2)^{10}$.
- So $P(X = 5) = \frac{\binom{10}{5}}{2^{10}}$

What if the coin is biased? Probability of H is 0.8?

A biased coin

You are throwing 10 biased coins with $P(\{H\}) = p$. What is the probability that the sum X equals 5?

- Probability of a length 10 binary sequence with 5 1's is
- So $P(X = 5) =$
- What about $P(X = 8)$?

A biased coin

You are throwing 10 biased coins with $P(\{H\}) = p$. What is the probability that the sum X equals 5?

- Probability of a length 10 binary sequence with 5 1's is $p^5(1-p)^5$ using independence.
- So $P(X = 5) =$
- What about $P(X = 8)$?

A biased coin

You are throwing 10 biased coins with $P(\{H\}) = p$. What is the probability that the sum X equals 5?

- Probability of a length 10 binary sequence with 5 1's is $p^5(1-p)^5$ using independence.
- So $P(X = 5) = \binom{10}{5} p^5(1-p)^5$.
- What about $P(X = 8)$?

A biased coin

You are throwing 10 biased coins with $P(\{H\}) = p$. What is the probability that the sum X equals 5?

- Probability of a length 10 binary sequence with 5 1's is $p^5(1-p)^5$ using independence.
- So $P(X = 5) = \binom{10}{5} p^5 (1-p)^5$.
- What about $P(X = 8)$? $\binom{10}{8} p^8 (1-p)^2$.

A biased coin

You are throwing 10 biased coins with $P(\{H\}) = p$. What is the probability that the sum X equals 5?

- Probability of a length 10 binary sequence with 5 1's is $p^5(1-p)^5$ using independence.
- So $P(X=5) = \binom{10}{5} p^5(1-p)^5$.
- What about $P(X=8)$? $\binom{10}{8} p^8(1-p)^2$.
- In general $P(X=k) = \binom{n}{k} p^k(1-p)^{n-k}$

The Binomial random variable

A biased coin ($P(\{H\}) = p$) is tossed n times. All tosses are mutually independent.

- Let X_i be a Bernoulli random variable which is 1 if the i^{th} toss gave a head. Then $\{X_i, i = 1, \ldots, n\}$ are independent Bernoullis.
- Let Y be the number of heads we see at the end of all n tosses.
- Whats the relationship of Y to the X_i's?

Y is called a **Binomial** random variable. What is the PMF of Y?

The Binomial random variable

A biased coin ($P(\{H\}) = p$) is tossed n times. All tosses are mutually independent.

- Let X_i be a Bernoulli random variable which is 1 if the i^{th} toss gave a head. Then $\{X_i, i = 1, \ldots, n\}$ are independent Bernoullis.
- Let Y be the number of heads we see at the end of all n tosses.
- Whats the relationship of Y to the X_i's?

Y is called a **Binomial** random variable. What is the PMF of Y?

- $P(Y = 0) = P(\{\text{no heads}\}) = (1 - p)^n$.

The Binomial random variable

A biased coin ($P(\{H\}) = p$) is tossed n times. All tosses are mutually independent.

- Let X_i be a Bernoulli random variable which is 1 if the i^{th} toss gave a head. Then $\{X_i, i = 1, \ldots, n\}$ are independent Bernoullis.
- Let Y be the number of heads we see at the end of all n tosses.
- Whats the relationship of Y to the X_i's?

Y is called a **Binomial** random variable. What is the PMF of Y?

- $P(Y = 0) = P(\{\text{no heads}\}) = (1 - p)^n$.
- $P(Y = n) = P(\{\text{all heads}\}) = p^n$.

The Binomial random variable

A biased coin ($P(\{H\}) = p$) is tossed n times. All tosses are mutually independent.

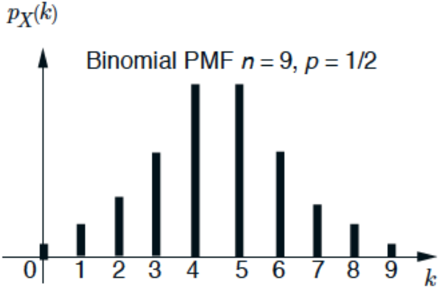
- Let X_i be a Bernoulli random variable which is 1 if the i^{th} toss gave a head. Then $\{X_i, i = 1, \ldots, n\}$ are independent Bernoullis.
- Let Y be the number of heads we see at the end of all n tosses.
- Whats the relationship of Y to the X_i's?

Y is called a **Binomial** random variable. What is the PMF of Y?

- $P(Y = 0) = P(\{\text{no heads}\}) = (1 - p)^n$.
- $P(Y = n) = P(\{\text{all heads}\}) = p^n$.
- $P(Y = k) = \binom{n}{k} p^k (1 - p)^{n-k}$ for $k \in \{0, 2, \ldots n\}$
- $\sum_{k=0}^{n} \binom{n}{k} p^k (1 - p)^{n-k} = 1$ (Binomial expansion)!

The Binomial random variable

- Sum of independent Bernoullis give a Binomial!
- We will denote the Binomial PMF by $Binomial(n, p)$ or $Bin(n, p)$.
- We will write $X \sim Binomial(n, p)$ to indicate that X is **distributed** as a Binomial random variable with parameters n and p.
- Quick look at the histogram of $X \sim Binomial(9, 1/2)$.



Urn example

You have a urn with an 10 blue and 20 red balls. You pick 9 balls with replacement. Let X be the number of blue balls. What is $P(\text{X}=3)$?

- With replacement is key here! I can draw the same ball twice!
- We have $P(\text{blue ball}) =$

Urn example

You have a urn with an 10 blue and 20 red balls. You pick 9 balls with replacement. Let X be the number of blue balls. What is $P(\mathrm{X}{=}3)$?

- With replacement is key here! I can draw the same ball twice!
- We have $P(\text{blue ball}) = 10/30 = 1/3$ and $P(\text{red ball}) =$

Urn example

You have a urn with an 10 blue and 20 red balls. You pick 9 balls with replacement. Let X be the number of blue balls. What is $P(\mathrm{X}{=}3)$?

- With replacement is key here! I can draw the same ball twice!
- We have $P(\text{blue ball}) = 10/30 = 1/3$ and $P(\text{red ball}) = 2/3$.

Urn example

You have a urn with an 10 blue and 20 red balls. You pick 9 balls with replacement. Let X be the number of blue balls. What is P(X=3)?

- With replacement is key here! I can draw the same ball twice!
- We have $P(\text{blue ball}) = 10/30 = 1/3$ and $P(\text{red ball}) = 2/3$.
- Probability of any sequence of 9 balls with 3 blue balls is
- How many sequences of 9 balls are there with 3 blue balls?

Urn example

You have a urn with an 10 blue and 20 red balls. You pick 9 balls with replacement. Let X be the number of blue balls. What is $P(\text{X}=3)$?

- With replacement is key here! I can draw the same ball twice!
- We have $P(\text{blue ball}) = 10/30 = 1/3$ and $P(\text{red ball}) = 2/3$.
- Probability of any sequence of 9 balls with 3 blue balls is $(1/3)^3(2/3)^6$.
- How many sequences of 9 balls are there with 3 blue balls?

Urn example

You have a urn with an 10 blue and 20 red balls. You pick 9 balls with replacement. Let X be the number of blue balls. What is $P(\mathrm{X}{=}3)$?

- With replacement is key here! I can draw the same ball twice!
- We have $P(\text{blue ball}) = 10/30 = 1/3$ and $P(\text{red ball}) = 2/3$.
- Probability of any sequence of 9 balls with 3 blue balls is $(1/3)^3(2/3)^6$.
- How many sequences of 9 balls are there with 3 blue balls? $\binom{9}{3}$ ways!

Urn example

You have a urn with an 10 blue and 20 red balls. You pick 9 balls with replacement. Let X be the number of blue balls. What is $P(\text{X}=3)$?

- With replacement is key here! I can draw the same ball twice!
- We have $P(\text{blue ball}) = 10/30 = 1/3$ and $P(\text{red ball}) = 2/3$.
- Probability of any sequence of 9 balls with 3 blue balls is $(1/3)^3(2/3)^6$.
- How many sequences of 9 balls are there with 3 blue balls? $\binom{9}{3}$ ways!
- So the probability $P(X = 3) = \begin{pmatrix} 9 \\ 3 \end{pmatrix}(1/3)^3(2/3)^6$!

Urn example

You have a urn with an 10 blue and 20 red balls. You pick 9 balls with replacement. Let X be the number of blue balls. What is P(X=3)?

- With replacement is key here! I can draw the same ball twice!
- We have $P(\text{blue ball}) = 10/30 = 1/3$ and $P(\text{red ball}) = 2/3$.
- Probability of any sequence of 9 balls with 3 blue balls is $(1/3)^3(2/3)^6$.
- How many sequences of 9 balls are there with 3 blue balls? $\binom{9}{3}$ ways!
- So the probability $P(X = 3) = \begin{pmatrix} 9 \\ 3 \end{pmatrix} (1/3)^3(2/3)^6$!
- X is a Binomial(9,3) random variable! $X \sim \mathit{Binomial}(9, 3)$

The Geometric random variable

- The Bernoulli PMF describes the probability of success/failure in a single trial.
- The Binomial PMF describes the probability of k successes out of n trials.
- Sometimes we may also be interested in doing trials until we see a success.

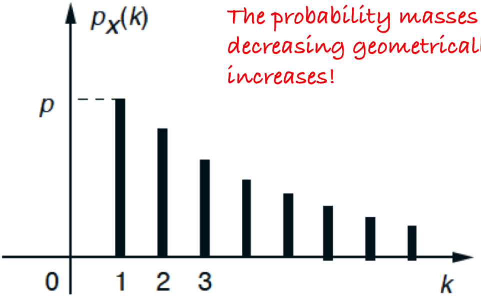
- Alice resolves to keep buying lottery tickets until he wins a hundred million dollars. She is interested in the random variable "number of lottery tickets bought until he wins the 100M$ lottery".
- Annie is trying to catch a taxi. How many occupied taxis will drive past before she finds one that is taking passengers?

- The number of trials required to get a single success is a **Geometric Random Variable**

The geometric random variable

We repeatedly toss a biased coin $(P(\{H\}) = p)$. The geometric random variable is the number X of tosses to get a head.

- X can take any integral value.
- $P(X = k) = P(\{\underbrace{TT\ldots T}_{k-1}H\}) = (1-p)^{k-1}p.$
- $\sum_k P(X = k) = 1$ (why?)



The probability masses are decreasing geometrically as k increases!

The Geometric random variable

- The Bernoulli PMF describes the probability of success/failure in a single trial.
- The Binomial PMF describes the probability of k successes out of n trials.
- Sometimes we may also be interested in doing trials until we see a success.

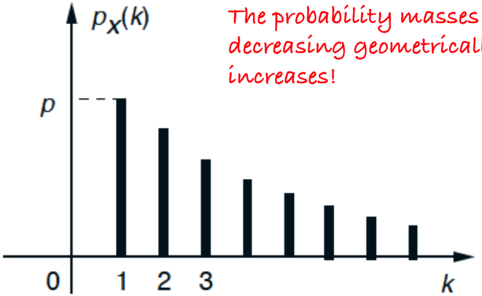
- Alice resolves to keep buying lottery tickets until he wins a hundred million dollars. She is interested in the random variable "number of lottery tickets bought until he wins the 100M$ lottery".
- Annie is trying to catch a taxi. How many occupied taxis will drive pass before she finds one that is taking passengers?

- The number of trials required to get a single success is a **Geometric Random Variable**

The geometric random variable

We repeatedly toss a biased coin ($P(\{H\}) = p$). The geometric random variable is the number X of tosses to get a head.

- X can take any integral value.
- $P(X = k) = P(\{\underbrace{TT \ldots T}_{k-1} H\}) = (1-p)^{k-1}p$.
- $\sum_k P(X = k) = 1$ (why?)



The probability masses are decreasing geometrically as k increases!

The geometric random variable

What is $P(X \geq k)$? What is $P(X > k)$?

The geometric random variable

What is $P(X \geq k)$? What is $P(X > k)$?

- $P(X \geq k) = \sum_{i=k}^{\infty} p(1-p)^{i-1} = (1-p)^{k-1}$

The geometric random variable

What is $P(X \geq k)$? What is $P(X > k)$?

- $P(X \geq k) = \sum_{i=k}^{\infty} p(1-p)^{i-1} = (1-p)^{k-1}$
- Intuitively, this is asking for the probability that the first $k-1$ tosses are tails.
- This probability is $P(X \geq k) = (1-p)^{k-1}$

The geometric random variable

What is $P(X \geq k)$? What is $P(X > k)$?

- $P(X \geq k) = \sum_{i=k}^{\infty} p(1-p)^{i-1} = (1-p)^{k-1}$
- Intuitively, this is asking for the probability that the first $k-1$ tosses are tails.
- This probability is $P(X \geq k) = (1-p)^{k-1}$
- $X > k$ is the event that $X \geq k+1$, and so $P(X > k) = (1-p)^{k}$

The memoryless property

What is $P(X > a + b | X > a)$?

The memoryless property

What is $P(X > a + b | X > a)$?

- $$P(X > a + b | X > a) = \frac{P(X > a + b)}{P(X > a)}$$
  $$= \frac{(1-p)^{a+b}}{(1-p)^a} = (1-p)^b$$
  $$= P(X > b)$$
- You forgot about $X > a$ and started the clock afresh!

The Poisson random variable

I have a book with 10000 words. Probability that a word has a typo is $1/1000$. I am interested in how many misprints can be there on average? So a Poisson often shows up when you have a Binomial random variable with very large n and very small p but $n \times p$ is moderate. Here $np = 10$.

Our random variable might be:

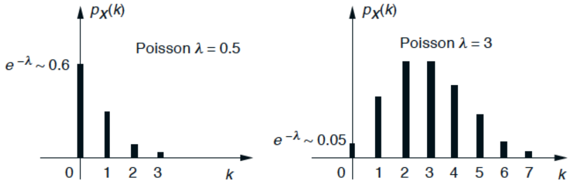
- The number of car crashes in a given day.
- The number of buses arriving within a given time period.
- The number of mutations on a strand of DNA.

We can describe such situations using a **Poisson random variable**.

The Poisson random variable

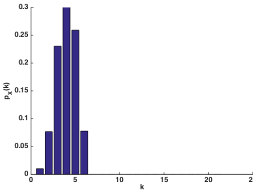
- A Poisson random variable takes non-negative integers as values. It has a nonnegative parameter λ.
- $P(X=k) = e^{-\lambda}\dfrac{\lambda^k}{k!}$, for $k = 0, 1, 2 \dots$.
- $\displaystyle\sum_{k=0}^{\infty} P(X=k) = e^{-\lambda}(1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \dots) = 1.$ (Exponential series!)

The PMF is monotonically decreasing for λ=0.5

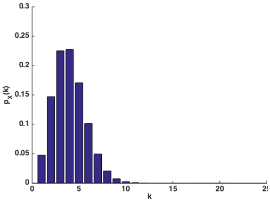


The PMF is increasing and then decreasing for λ=3

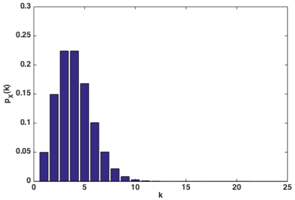
Poisson random variable



Binomial(5,0.6)



Binomial(100,0.03)



Poisson(3)

- When n is very large and p is very small, a binomial random variable can be well approximated by a Poisson with $\lambda = np$.
- In the above figure we increased n and decreased p so that $np = 3$.
- See how close the PMF's of the Binomial(100,0.03) and Poisson(3) are!
- More formally, we see that $\binom{n}{k} p^k (1-p)^{n-k} \approx \frac{e^{-\lambda}\lambda^k}{k!}$ when n is large, k is fixed, and p is small and $\lambda = np$.

Example

Assume that on a given day 1000 cars are out in Austin. On an average three out of 1000 cars run into a traffic accident per day.

1. What is the probability that we see at least two accidents in a day?

4. If you know there is at least one accident, what is the probability that the total number of accidents is at least two?

Example

Assume that on a given day 1000 cars are out in Austin. On an average three out of 1000 cars run into a traffic accident per day.

1. What is the probability that we see at least two accidents in a day?

2. Use poisson approximation!

4. If you know there is at least one accident, what is the probability that the total number of accidents is at least two?

Example

Assume that on a given day 1000 cars are out in Austin. On an average three out of 1000 cars run into a traffic accident per day.

1. What is the probability that we see at least two accidents in a day?

2. Use poisson approximation!
3. $P(X \geq 2) = 1 - P(X = 0) - P(X = 1) = 1 - e^{-3}(1 + 3) = 0.8$

4. If you know there is at least one accident, what is the probability that the total number of accidents is at least two?

Example

Assume that on a given day 1000 cars are out in Austin. On an average three out of 1000 cars run into a traffic accident per day.

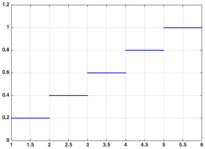
1. What is the probability that we see at least two accidents in a day?
2. Use poisson approximation!
3. $P(X \geq 2) = 1 - P(X = 0) - P(X = 1) = 1 - e^{-3}(1 + 3) = 0.8$
4. If you know there is at least one accident, what is the probability that the total number of accidents is at least two?
5. $P(X \geq 1) = 1 - P(X = 0) = 1 - e^{-3} = 0.950.$
   $P(X \geq 2 | X \geq 1) = P(X \geq 2)/P(X \geq 1) = 0.8/0.950 = 0.84$

Cumulative Distribution Functions

- For any random variable the cumulative distribution function is defined as:

$$F_X(a) = \sum_{x \leq a} p(x)$$

- Can you work out the PMF of the following random variable?



Function of a random variable

- A function of a random variable is also a random variable.
- Let X be the number of heads in 5 fair coin tosses.
- We know that X has the Binomial(5,1/2) distribution.
- Define $Y = X \mod 4$. Whats its PMF?

Function of a random variable

- A function of a random variable is also a random variable.
- Let X be the number of heads in 5 fair coin tosses.
- We know that X has the Binomial(5,1/2) distribution.
- Define $Y = X \mod 4$. Whats its PMF?
- $Y = \begin{cases} 0 & X \in \{0, 4\} \\ 1 & X \in \{1, 5\} \\ 2 & X = 2 \\ 3 & X = 3 \end{cases}$

Function of a random variable

- A function of a random variable is also a random variable.
- Let X be the number of heads in 5 fair coin tosses.
- We know that X has the Binomial(5,1/2) distribution.
- Define $Y = X \mod 4$. Whats its PMF?
- $Y = \begin{cases} 0 & X \in \{0,4\} \\ 1 & X \in \{1,5\} \\ 2 & X = 2 \\ 3 & X = 3 \end{cases}$
- Lets write down the PMF of Y.
  - $P(Y=0) = P(X=0) + P(X=4) = (1/2)^5 + \binom{5}{4}(1/2)^5$
  - $P(Y=1) = P(X=1) + P(X=5)$...and so on.

Function of a random variable

- A function of a random variable is also a random variable.
- Let X be the number of heads in 5 fair coin tosses.
- We know that X has the Binomial(5,1/2) distribution.
- Define $Y = X \mod 4$. Whats its PMF?
- $Y = \begin{cases} 0 & X \in \{0, 4\} \\ 1 & X \in \{1, 5\} \\ 2 & X = 2 \\ 3 & X = 3 \end{cases}$
- Lets write down the PMF of Y.
  - $P(Y = 0) = P(X = 0) + P(X = 4) = (1/2)^5 + \binom{5}{4}(1/2)^5$
  - $P(Y = 1) = P(X = 1) + P(X = 5)$...and so on.
  - More formally, if $Y = g(X)$ then we have:

$$p_Y(y) = \sum_{\{x|g(x)=y\}} p_X(x).$$