Chapter 4

Discrete Random Variables

Yingbo Li

Clemson University

MATH 4000 / 6000

Recap

Independence

- $P(A \cap B) = P(A) \times P(B)$, $P(A|B) = P(A)$
- Events $A_1, A_2, \ldots, A_n$ is (mutually) independent if for any $1 \le r \le n$ of them

$$P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_r}) = P(A_{i_1})P(A_{i_2}) \cdots P(A_{i_r})$$

Question

Which of the following statements is false?

(a) Two disjoint events cannot occur at the same time.

(b) *Two independent events cannot occur at the same time.*

(c) Two complementary events cannot occur at the same time.

Question

The figure and table on the right show the average daily probabilities for being born on any day in a certain month. If we randomly select two people, what is the probability that the first one is born in April and the second in July?

Jan	0.0026123	*Jul*	*0.0028655*
Feb	0.0026785	Aug	0.0028954
Mar	0.0026838	Sep	0.0029407
Apr	*0.0026426*	Oct	0.0027705
May	0.0026702	Nov	0.0026842
Jun	0.0027424	Dec	0.0026864

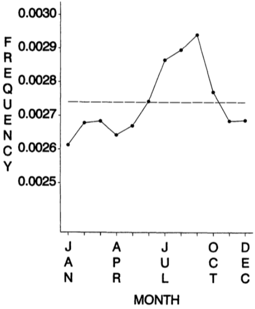


Figure 1. Average Daily United States Birth Frequencies, by Month. The rate increases fairly steadily from about 5% below the average in January to about 7% above it in September. The dashed line shows the uniform birth probability distribution of 1/365.

(a) $0.0026426 \times 0.0028655$

(b) $(0.0026426 \times 30) + (0.0028655 \times 31)$

(c) $(0.0026426 \times 30) \times (0.0028655 \times 31)$

(d) $(0.0026426^{30}) \times (0.0028655^{31})$

Nunnikhoven (1992). *A Birthday Problem Solution for Nonuniform Birth Frequencies*.

Outline

Random Variables

Definition

Random Variable X *is a real-valued function on the sample space* S*.*

Example: If $S = \{a = (a_1, a_2) : 1 \leq a_1, a_2 \leq 6\}$ is the 36 element space resulting from rolling two fair six-sided dice, then the following are all random variables

$$X(a) = a_1$$
$$Y(a) = |a_1 - a_2|$$
$$Z(a) = a_1 + a_2$$

- Random variable is a number associated with a random experiment.
- Random variables are in essence a fancy way of describing an event, e.g.

$$P(X = 1) = 1/6$$

Example: a coin is flipped until a head is obtained. The flips are Independent and each one has probability p of heads. Let the random variable X denote the number of flips until a head is obtained.

$$
\begin{aligned}
P(X=1) &= P(\mathsf{H}) = p \\
P(X=2) &= P(\mathsf{TH}) = (1-p)p \\
P(X=3) &= P(\mathsf{TTH}) = (1-p)^2 p \\
&\cdots\cdots \\
P(X=n) &= P(\mathsf{TT}\cdots\mathsf{TH}) = (1-p)^{n-1}p\cdots \qquad \cdots
\end{aligned}
$$

Cumulative distribution function (cdf)

Definition

For a random variable X, the function F defined by

$$F(x) = P(X \leq x), \quad -\infty < x < \infty$$

is called the cumulative distribution function of X.

Note that

- Capital X: random variable
- Little x: a real-valued number
- $\leq$: smaller than or equal to

In the previous coin flipping example,

$$F(n) = P(X \leq n) = \sum_{i=1}^{n} (1-p)^{i-1} p = \frac{[1-(1-p)^n]p}{1-(1-p)} = 1-(1-p)^n$$

Properties of the cdf

- $P(a < X \leq b) = F(b) - F(a)$, for all $a < b$
- $P(X < b)$ does not necessary equal to $P(X \leq b)$.

1. $F(x)$ is a non-decreasing function; i.e., if $x_1 < x_2$, then
$$F(x_1) \leq F(x_2)$$
2. $$\lim_{x\to\infty} F(x) = 1 \iff P(X \leq \infty) = 1$$
3. $$\lim_{x\to-\infty} F(x) = 0 \iff P(X > -\infty) = 1$$
4. $F(x)$ is right continuous; i.e., for any decreasing sequence $\{x_n : n = 1, 2, \ldots\}$ that converges to x,
$$\lim_{n\to\infty} F(x_n) = F(x) \Longrightarrow \lim_{n\to\infty} P(X \leq x + \frac{1}{n}) = P(X \leq x)$$

The cdf tells us everything about a rv X

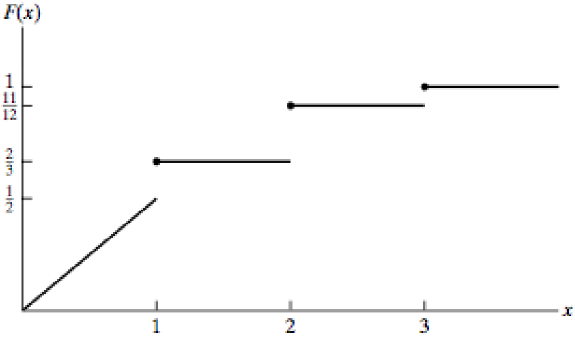


FIGURE 4.8: Graph of $F(x)$.

- $P(X \le 0.5) = 0.5 \times \frac{1/2}{1} = 0.25$
- $P(X > 2) = 1 - P(X \le 2) = \frac{1}{12}$
- $P(X < 1) = 0.5, P(X \le 1) = \frac{2}{3}$
- $P(X = 1) = P(X \le 1)$ $-P(X < 1) = \frac{2}{3} - \frac{1}{2} = \frac{1}{6}$

Discrete random variables

Definition

- *A random variable X that can take on at most a countable number of possible values is a discrete random variable.*
- *For a discrete random variable X, we define the probability mass function (pmf) by*

$$p(x) = P(X = x)$$

Note: pmf is also written as $f(x)$.

Examples of discrete rv

- Flip a coin 10 times, $X =$ number of heads.
  X can be $0, 1, 2, \ldots, 10$, finite. $\Longrightarrow$ discrete rv.
- $X =$ number of your ex boyfriends / girlfriends.
  X can be $0, 1, 2, \ldots, \infty$, countable. $\Longrightarrow$ discrete rv.
- $X =$ a random number in $[0, 1]$ generated by computer
  X can anything in $[0, 1]$, uncountable. $\Longrightarrow$ not discrete rv.

pmf and cdf

- For a discrete rv X, there exists a countable sequence $x_1, x_2, \ldots$, st

$$\begin{aligned} p(x_i) &> 0 \quad \text{for } i = 1, 2, \ldots \\ p(x) &= 0 \quad \text{for all other values of } x \end{aligned}$$

  and

$$\sum_{i=1}^{\infty} p(x_i) = 1$$

- Relationship between pmf and cdf (for discrete rv)

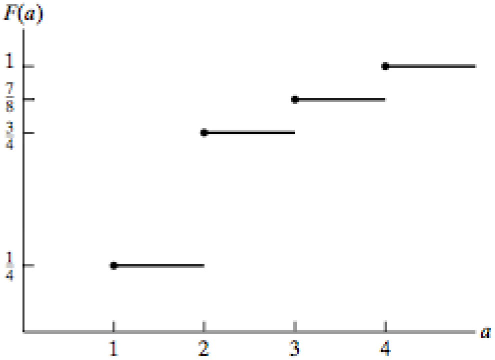
$$F(a) = \sum_{\text{all } x \leq a} p(x)$$

- If we know pmf, we can compute cdf. And vice versa.

X has a pmf given by

$$p(1) = \frac{1}{4},\ p(2) = \frac{1}{2},\ p(3) = \frac{1}{8},\ p(4) = \frac{1}{8}$$

$$F(a) = \begin{cases} 0 & a < 1 \\ \frac{1}{4} & 1 \le a < 2 \\ \frac{3}{4} & 2 \le a < 3 \\ \frac{7}{8} & 3 \le a < 4 \\ 1 & 4 \le a \end{cases}$$



Recap

Random Variable

- Random Variable X is a real-valued function on the sample space S.

$$X : S \longrightarrow \mathbb{R}$$

- Cumulative distribution function (cdf)

$$F_X(x) = P(X \leq x), \quad \text{for any } x \in \mathbb{R}$$

Discrete random variable

- can only take at most a countable number of possible values.
- probability mass function (pmf)

$$p_X(x) = P(X = x)$$

- $\sum_{i=1}^{\infty} p(x_i) = 1$
- $F(a) = \sum_{\text{all } x \leq a} p(x)$

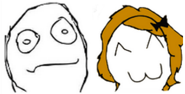
Expected value

Definition

The expected value (or mean) of a discrete random variable is defined as

$$E[X] = \sum_{x:p(x)>0} x \cdot P(X = x) = \sum_{x:p(x)>0} xp(x)$$

- $E[X]$ is a weighted average of the possible values x that X can take on, each value being weighted by the probability $p(x)$.



When she told me I was average,
she was just being mean.

Example: X is the outcome when we roll a 4-sided fair die. Find $E[X]$.

$$p(1) = p(2) = p(3) = p(4) = \frac{1}{4}$$

$$E[X] = 1 \times \frac{1}{4} + 2 \times \frac{1}{4} + 3 \times \frac{1}{4} + 4 \times \frac{1}{4} = 2.5$$

Toss a coin. Suppose the probability of a head is p. Let X be a 0-1 indicator random variable s.t.

$$X = \begin{cases} 1 & \text{if head is obtained} \\ 0 & \text{otherwise} \end{cases}$$

Compute $\mu = E[X]$.

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

- In general, for indicator variable $X = \delta_A$ or $\mathbf{1}_A$, denoted as

$$X = \begin{cases} 1 & \text{if event } A \text{ occurs} \\ 0 & \text{otherwise} \end{cases}$$

The expected value of X equals the probability that A occurs.

$$E[X] = 1 \cdot P(A) + 0 \cdot P(A^c) = P(A)$$

Question

Let the rv X denote the GP a certain student will earn in this class. Suppose its pmf is

$$p(0) = 0.05, \quad p(1) = 0.05, \quad p(2) = 0.3, \quad p(3) = 0.4$$

Calculate his/her expected GP $E[X]$.

Since the domain of X (all possible values X can take) is $\{0, 1, 2, 3, 4\}$, first compute $p(4)$.

$$p(4) = 1 - p(0) - p(1) - p(2) - p(3) = 0.2$$

Then compute $E[X]$

$$E[X] = 0 \times 0.05 + 1 \times 0.05 + 2 \times 0.3 + 3 \times 0.4 + 4 \times 0.2 = 2.65$$

Expectation of a function of a random variable

- If X is a discrete rv, and g is a real-valued function then the expectation (or expected value) of $Y = g(X)$ is

$$E[g(X)] = \sum_{x:p_X(x)>0} g(x) \cdot p_X(x)$$

X is a discrete rv with pmf

x	-1	0	1
$p_X(x)$	$1/4$	$1/2$	$1/4$

$Y = X^2$. Compute $E[Y]$.

$$\begin{aligned} E[Y] &= \sum_{\text{all } x} x^2 \cdot p_X(x) \\ &= (-1)^2 \times \frac{1}{4} + 0^2 \times \frac{1}{2} + 1^2 \times \frac{1}{4} \\ &= \frac{1}{2} \end{aligned}$$

Properties of expected values

If a and b are constants, then

$$E[aX + b] = aE[X] + b$$

- Holds for all rv X (not necessary discrete rv).
- Proof for discrete rv: let $g(X) = aX + b$

Special cases of linear transformation $E[aX + b] = aE[X] + b$

- constant factor

$$E[aX] = aE[X]$$

- constant

$$E[b] = b$$

Recap

Expectation μ

- **For discrete rv** - $E[X] = \sum_{\text{all } x} x \cdot p(x)$
- **Functions** - $E[g(X)] = \sum_{\text{all } x} g(x)\ p(x)$
- **Indicators** - $E[\delta_A] = P(A)$ where δ_A is an indicator function
- **Linear function** - $E[aX + b] = aE[X] + b$
- **Constants** - $E[c] = c$ if c is constant

For two rv's X and Y

$$E[X + Y] = E[X] + E[Y]$$

Variance

- Expected value (or mean) $\mu = E[X]$ yields the weighted average of the possible values of X

Definition

Variance measures the variation (or spread) of these values

$$\sigma^2 = Var(X) = E\left[(X - E(X))^2\right] = E\left[(X - \mu)^2\right]$$

This holds for all rv X (not necessary discrete rv).

- One common simplification:

$$Var(X) = E(X^2) - \mu^2$$

Standard deviation

Definition

Standard Deviation is the square root of the variance

$$\sigma = SD(X) = \sqrt{Var(X)}$$

Example: compute the standard deviation of a fair 4-sided die toss.

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

$$E[X] = 1 \times \frac{1}{4} + 2 \times \frac{1}{4} + 3 \times \frac{1}{4} + 4 \times \frac{1}{4} = 2.5$$

$$E[X^2] = 1 \times \frac{1}{4} + 4 \times \frac{1}{4} + 9 \times \frac{1}{4} + 16 \times \frac{1}{4} = 7.5$$

$$Var(X) = 7.5 - 2.5^2 = 1.25$$

$$SD(X) = \sqrt{Var[X]} = 1.12$$

Variance measures the spread of X

X_1 is a discrete rv with pmf

x	-1	1
$p_{X_1}(x)$	$1/2$	$1/2$

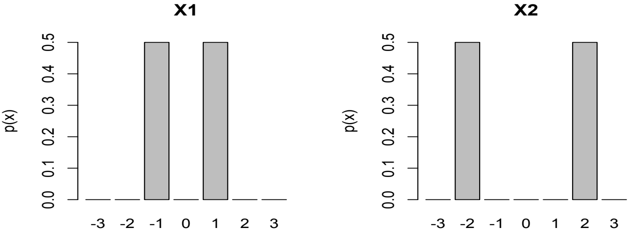
X_2 is a discrete rv with pmf

x	-2	2
$p_{X_2}(x)$	$1/2$	$1/2$

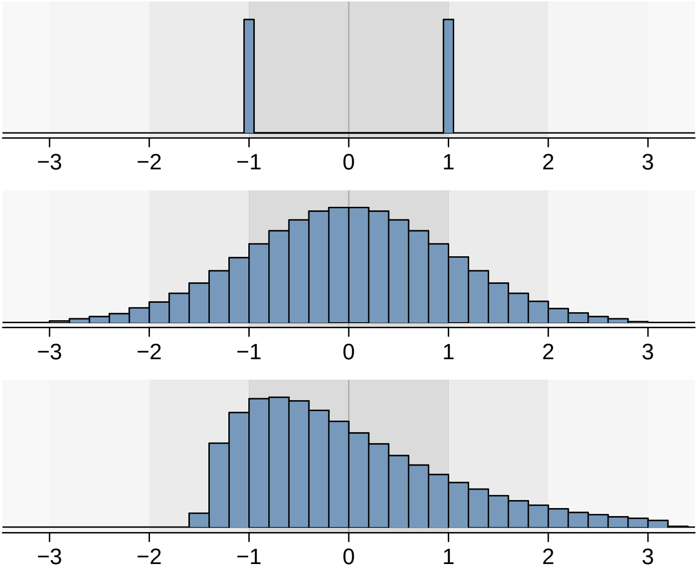
$$\mu_1 = 0,\ \sigma_1^2 = E[X_1^2] - 0^2 = 1,\ \sigma_1 = 1$$

$$\mu_2 = 0,\ \sigma_2^2 = E[X_2^2] - 0^2 = 4,\ \sigma_2 = 2$$

Increasing variance (or sd) reflects increasing spread.



Distributions with SD = 1



Property of variance

Question

We know that $Var(X) = E[(X-\mu)^2]$, and for constants a, b,

$$E[aX+b] = aE[X] + b.$$

Write $Var(aX+b)$ as a function of $Var(X)$.

$$\begin{aligned} Var(aX+b) &= E[(aX+b-E[aX+b])^2] \\ &= E[(aX+b-aE[X]-b)^2] \\ &= E[(aX-a\mu)^2] \\ &= E[a^2(X-\mu)^2] \\ &= a^2 Var(X) \end{aligned}$$

Properties of variance

$Var(X) \geq 0$

- $Var(X) = 0$ if and only if X is a constant.

If a and b are constants, then

$$Var(aX + b) = a^2 Var(X)$$

- $$Var(aX) = a^2 Var(X)$$
- $$Var(X + b) = Var(X)$$
- $$Var(b) = 0$$

Bernoulli distribution

A trial has two outcomes: success (1) or failure (0).
Let rv X be the number of success in a single trial.

Definition

Random variable X has a Bernoulli distribution, if

$$P(X = 1) = p, \quad P(X = 0) = 1 - p$$

where $0 \le p \le 1$ is the probability of a success.

- The pmf of Bernoulli distribution

$$X \sim \text{Ber}(p) \iff p(1) = p, \; p(0) = 1 - p$$

- Found by Switzerland mathematician Jacob Bernoulli.

Examples of Bernoulli distributions

- Toss a fair coin and obtain a head. $p = 0.5$.
- Roll a fair 6-sided die and obtain a 6. $p = 1/6$.
- Earn an A for this class. $p \in [0, 1]$.

Question

Find the mean and variance of $X \sim \text{Ber}(p)$.

$$E(X) = 1 \times p + 0 \times (1 - p) = p$$

$$Var(X) = E(X^2) - (E[X])^2 = 1^2 \times p + 0^2 \times (1 - p) - p^2 = p - p^2 = p(1 - p)$$

Recap

Variance σ^2

- **For all rv**

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

- **Linear function**

$$Var(aX+b) = a^2 Var(X)$$

- **Constants**

$$Var(c) = 0$$

Standard deviation

$$SD(X) = \sqrt{Var(X)}$$

Bernoulli distribution

$$p(1) = p, \quad p(0) = 1-p$$

$$\mu = p, \quad \sigma^2 = p(1-p)$$

Binomial distribution

Definition

Define X to be the number of successes in a fixed number n of independent trials with the same probability of success p as having a Binomial distribution.

Then the pmf of X is $P(X=k) = P($getting k successes in n trials$)$

$$X \sim Bin(n,p) \iff p(k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \ldots, n$$

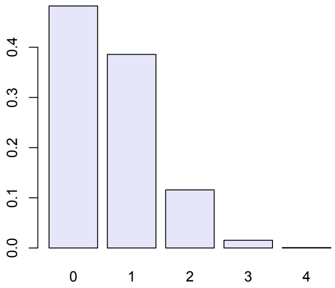
Example of Binomial distributions

- The number of heads obtained by tossing a fair coin 10 times. $n = 10, p = 0.5.$
- The number of 6 obtained when roll four fair 6-sided dice simultaneously. $n = 4, p = 1/6.$
- The number of A's students will earn in this semester. $n =$ number of classes you're taking, p varies by classes... $\Longrightarrow$ not really a Binomial rv!

Example: let X be the number of 6 obtained when roll four fair 6-sided dice simultaneously. Find its pmf.

$$n = 4,\ p = 1/6,$$

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



$$p(0) = \binom{4}{0} (1/6)^0 (5/6)^4 = 0.4823$$

$$p(1) = \binom{4}{1} (1/6)^1 (5/6)^3 = 0.3858$$

$$p(2) = \binom{4}{2} (1/6)^2 (5/6)^2 = 0.1157$$

$$p(3) = \binom{4}{3} (1/6)^3 (5/6)^1 = 0.0154$$

$$p(4) = \binom{4}{4} (1/6)^4 (5/6)^0 = 0.0008$$

Question

Example : let X be the number of 6 obtained when roll four fair 6-sided dice simultaneously. Its pmf is.

x	0	1	2	3	4
$p(x)$	0.4823	0.3858	0.1157	0.0154	0.0008

Find its cdf $F(x)$.

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ 0.4823 & \text{if } 0 \leq x < 1 \\ 0.8681 & \text{if } 1 \leq x < 2 \\ 0.9838 & \text{if } 2 \leq x < 3 \\ 0.9992 & \text{if } 3 \leq x < 4 \\ 1 & \text{if } x \geq 4 \end{cases}$$

Question

Suppose 42% of the students in a class have part time jobs. Among a random sample of 10 students in that class, what is the probability that exactly 8 have part time jobs?

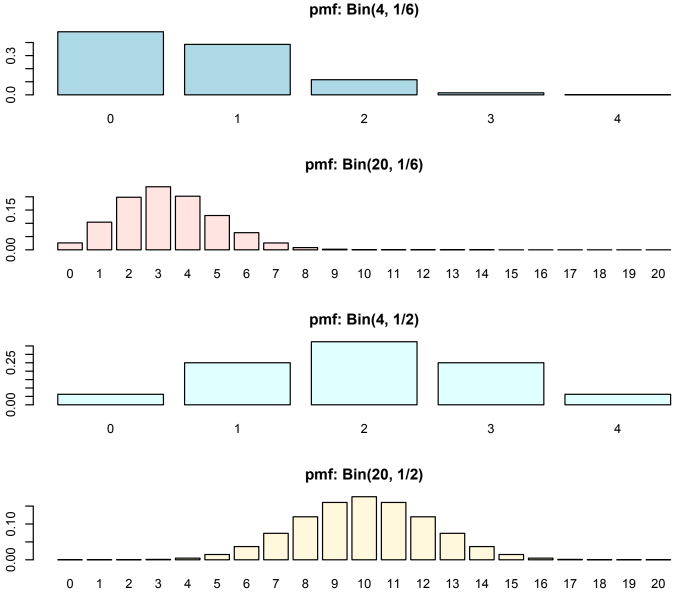
Binomial distribution. $n = 10, p = 0.42$.

$$P(X = 8) = \binom{10}{8} \times 0.42^8 \times 0.58^2 = 0.0147$$

We can use the binomial distribution to calculate the probability of k successes in n trials, as long as

1. the trials are independent
2. the number of trials, n, is fixed
3. each trial outcome can be classified as a *success* or a *failure*
4. the probability of success, p, is the same for each trial

Pmf of Binomial distribution: uni-modal



Binomial pmf is valid (or well-defined)

- Positive: for any $x \in \mathbb{R}$
$$p(x) \geq 0$$
- Total one:
$$\sum_{k=0}^{n} p(k) = 1$$

Recall the Binomial Theorem $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$

Mean of Binomial random variable

Toss a coin n times, each toss has prob p being a head. On average, total number of heads equals np.

$$E[X] = np$$

Properties of Binomial distribution

- Variance

$$Var[X] = np(1-p)$$

- If we have independent Bernoulli rv's $X_1, X_2, \dots, X_n$ with the same probability of success p, then their sum has a Binomial distribution

$$X = X_1 + X_2 + \cdots + X_n \sim \text{Bin}(n, p)$$

Question

Roll 4 fair six-sided dice. Let X be the number of 6 obtained. Find its mean and variance.

$$X \sim \mathsf{Bin}(n = 4, p = 1/6)$$

$$E[X] = np = 4 \times (1/6) = 2/3$$
$$Var(X) = np(1 - p) = 4 \times (1/6) \times (5/6) = 5/9$$

Recap

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

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

- mean $\mu = np$
- variance $\sigma^2 = np(1-p)$

Question

Random variable $X \sim \text{Binom}(n, p)$, and the value of n is fixed. For any fixed n, find p such that the distribution of X has the largest spread.

Variance (or standard deviation) is a measurement of spread.

$$\sigma^2 = np(1-p)$$

To find the $\hat{p}$ that maximize the spread, i.e., find the root of the derivative

$$\frac{d\sigma^2}{dp} = n(1-p+(-p)) = 0$$
$$\implies \hat{p} = 1/2$$

Count the number of ...

In beer brewing, cultures of yeast are kept alive in jars of fluid before being put into the mash.

- It's critical to control the amount of yeast used.
- Number of yeast cells in a fluid sample can be seen under a microscopes.
- Yeast cells are constantly multiplying and dividing.
- A famous statistician, Wiliam Sealy Gosset (aka "Student"), who worked for the Guinness Brewing Compnay in early 1900's, modeled the counts of yeast cells using the *Poisson distribution*.

The Poisson distribution was used in 1898 to count the number of soldiers in the Prussian Army who died accidentally from horse kicks

Poisson distribution

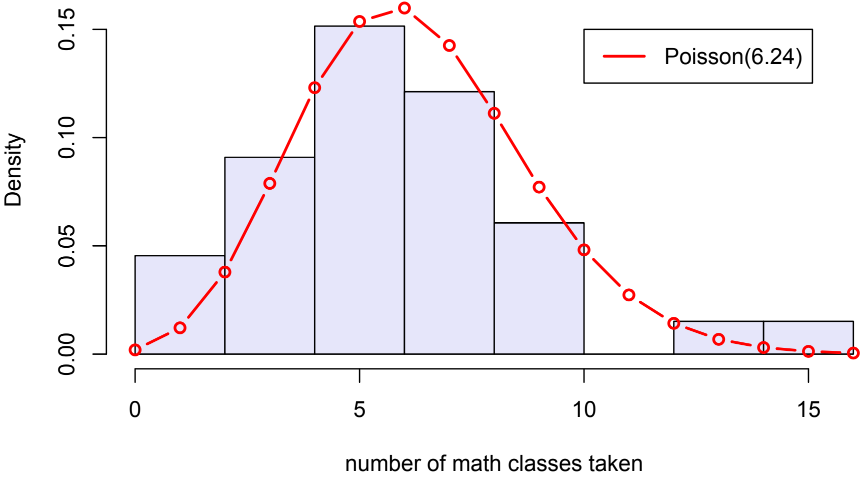
Definition

Denote rv X that takes value in $\{0, 1, 2, \ldots\}$ as having a Poisson distribution with parameter λ if its pmf is

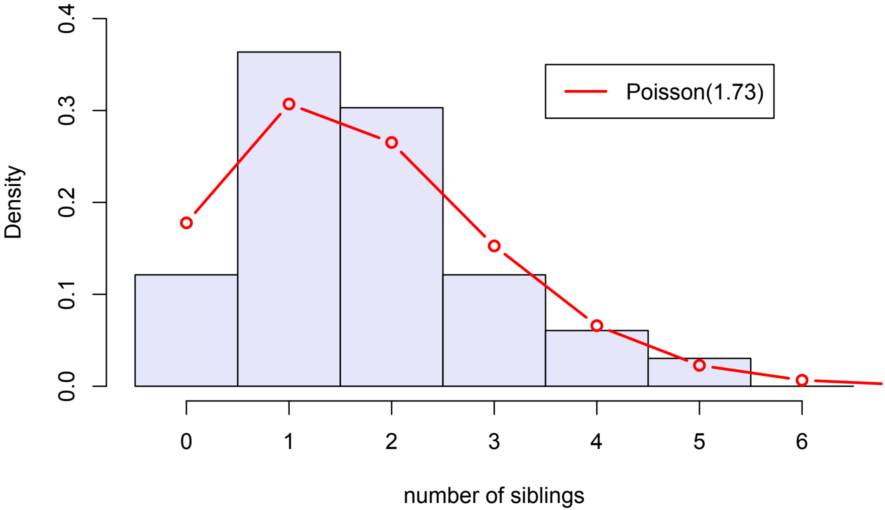
$$X \sim P(\lambda) \iff p(k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \ldots$$

- Formulated by French mathematician Siméon Denis Poisson
- Usually is used to model "the number of xxx occur". Hence lower bound is 0, no upper bound.
- Examples
  - Number of rains in this year
  - Number of mis-placed books in the Cooper library
  - Number of roses you will receive on the next Valentines Day

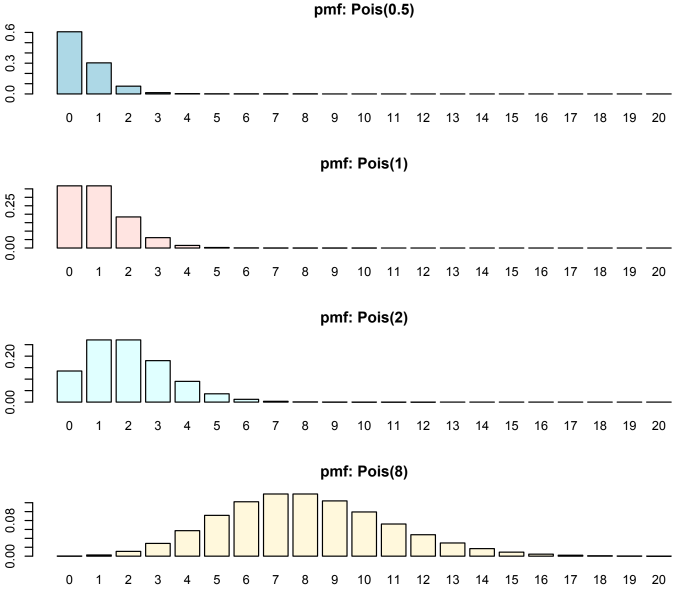
Example: number of Math classes taken



Example: number of siblings



Pmf of Poisson distribution



Properties of Poisson distribution $p(k) = e^{-\lambda} \frac{\lambda^k}{k!}$

- Well-defined (validness of pmf): non-negative, and

$$\sum_{k=0}^{\infty} p(k) = 1$$

- Taylor Series

$$f(x) = f(x_0) + \frac{x - x_0}{1!} f'(x_0) + \frac{(x - x_0)^2}{2!} f''(x_0) + \cdots$$

- Use Taylor series to verify that the Poisson distribution is well-defined

$$e^{\lambda} = 1 + \lambda + \frac{\lambda^2}{2!} + \cdots + \frac{\lambda^k}{k!} + \cdots$$

Properties of Poisson distribution $p(k) = e^{-\lambda} \frac{\lambda^k}{k!}$

- Mean

$$E[X] = \lambda$$

Properties of Poisson distribution $p(k) = e^{-\lambda} \frac{\lambda^k}{k!}$

- Variance

$$Var[X] = \lambda$$

Question

Suppose Clemson students have 1.73 siblings on average. What's the probability that a randomly selected student has at most one sibling?

$$X \sim \mathsf{P}(\lambda), E[X] = 1.73 \Longrightarrow \lambda = 1.73$$

$$\begin{aligned} P(X \leq 1) &= p(0) + p(1) \\ &= e^{-\lambda}\frac{\lambda^0}{0!} + e^{-\lambda}\frac{\lambda^1}{1!} \\ &= 0.1773 + 0.3067 = 0.4840 \end{aligned}$$

Example: find the probability that a randomly selected student has odd number of siblings.

$$X \sim \mathsf{P}(\lambda), \lambda = 1.73$$

$$\begin{aligned} P(X \text{ is an odd number}) &= p(1) + p(3) + p(5) + \cdots \\ &= e^{-\lambda}\frac{\lambda^1}{1!} + e^{-\lambda}\frac{\lambda^3}{3!} + e^{-\lambda}\frac{\lambda^5}{5!} + \cdots \end{aligned}$$

Note that

$$\begin{aligned} e^x &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots \\ e^{-x} &= 1 - x + \frac{x^2}{2!} - \frac{x^3}{3!} + \frac{x^4}{4!} - \cdots \end{aligned}$$

$$P(X \text{ is an odd number}) = e^{-\lambda}\left(\frac{e^{\lambda} - e^{-\lambda}}{2}\right) = \frac{1 - e^{-2\lambda}}{2} = 0.4843$$

Use Poisson to approximate Binomial distribution

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

- p: small
- n: large
- $\lambda = np$: of moderate size

then the distribution of X can be approximated by $\text{P}(\lambda)$.

Recap

Poisson distribution $X \sim \mathsf{P}(\lambda)$

$$p(k) = e^{-\lambda}\ \frac{\lambda^k}{k!}$$

- mean $\mu = \lambda$
- variance $\sigma^2 = \lambda$
- Approximate Binomial distribution with small p, large n, moderate np

$$\mathsf{P}(np) \approx \mathsf{Bin}(n, p)$$

Question

$X \sim \mathrm{P}(\lambda)$. Which is the following is FALSE?

(a) The mean and standard deviation of X are different.

(b) Pmf of X can be a decreasing function.

(c) *λ can only take values $0, 1, 2, \ldots$.*

(d) None of the above.

Geometric distribution

A gambler plays at a roulette table and alway bet on red until he wins...
In each round, his chance of winning is $18/38 = 0.47$.
Let X denote the number of rounds he plays.

Definition

Denote rv X that takes value in $\{1, 2, \ldots\}$ as having a Geometric distribution with parameter $p \in (0, 1)$ if its pmf is

$$X \sim \text{Geometric}(p) \iff p(k) = (1-p)^{k-1} p, \quad k = 1, 2, \ldots$$

- X represents the number of trials performed until we get a success, where p is the probability of success on each trial.
- Note the difference between Geometric distribution and Binomial distribution! Eg: whether the total number of trials is fixed.

Properties of Geometric distribution $p(k) = (1-p)^{k-1}p$

- Well-defined (validness of pmf): non-negative, $\sum_{k=1}^{\infty} p(k) = 1$

$$\sum_{k=1}^{\infty}(1-p)^{k-1}p = p[1 + (1-p) + (1-p)^2 + \cdots] = \frac{p}{1-(1-p)} = 1$$

- Cdf: $P(X \leq k) = 1 - (1-p)^k$

$$P(X \geq k+1) = P(\text{The first } k \text{ trials all fail})$$

- Mean

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

- Variance

$$Var[X] = \frac{1-p}{p^2}$$

Gambler's fallacy

If the gambler loses 5 times in a row, will he more likely to win in the 6th round?

$$P(X > 6 \mid X > 5) \overset{?}{<} P(X > 1)$$

Unfortunately, not.

Definition

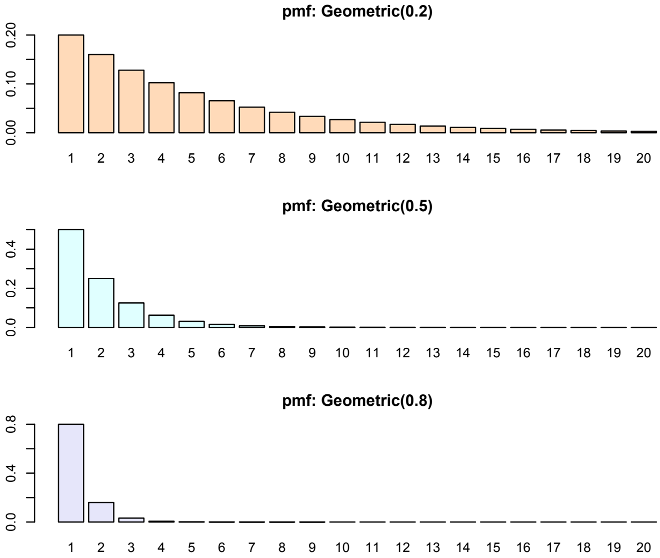
We say a distribution is memoryless, if

$$P(X > n + k \mid X > n) = P(X > k)$$

- Geometric rv is memoryless.

$$\begin{aligned} P(X > n + k \mid X > n) &= \frac{P(X > n + k)}{P(X > n)} \\ &= \frac{(1-p)^{n+k}}{(1-p)^n} = (1-p)^k = P(X > k) \end{aligned}$$

Pmf of Geometric distribution



Example: X denotes the number of times a die is rolled until 6 is obtained.

1. What are the odds we have to roll it 10 or more times?
2. How many times do we expect to roll?
3. Find $Var[X]$.

1. $X \sim$ Geometric$(p = 1/6)$, and

$$P(X \geq 10) = 1 - P(X \leq 9) = 1 - [1 - (1-p)^9] = (5/6)^9$$

2. Mean of Geometric distribution

$$E[X] = 1/p = 6$$

3. Variance

$$Var[X] = \frac{1-p}{p^2} = \frac{\frac{5}{6}}{\frac{1}{6} \cdot \frac{1}{6}} = 30$$

Negative Binomial distribution

Definition

Denote rv X that takes value in $\{1, 2, \ldots\}$ as having a Negative Binomial distribution with parameter $p \in (0,1)$ if its pmf is

$$X \sim NB(r,p) \iff p(k) = \binom{k-1}{r-1}(1-p)^{k-r}p^r, \quad k = r, r+1, \ldots$$

- X represents the number of trials performed until we get r success, where p is the probability of success on each trial.
- Well-defined (validiness of pmf).
- Connection between Negative Binomial and Geometric distributions

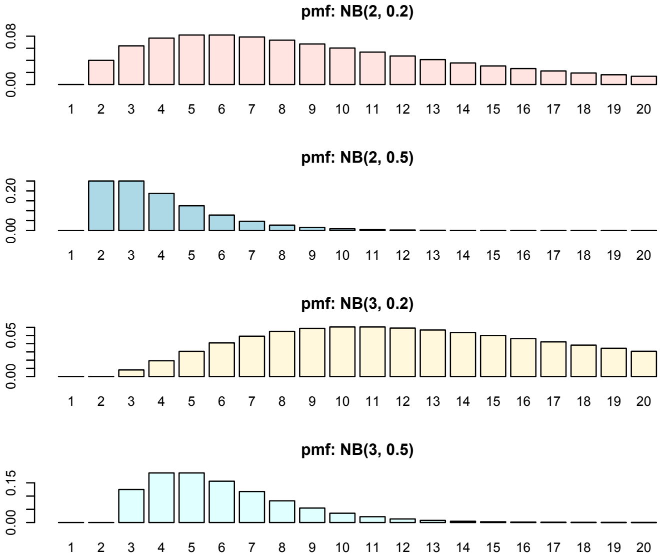
$$X \sim \text{Geometric}(p) \iff X \sim \text{NB}(1,p)$$

Note: Proof of $\sum_{k=r}^{\infty} p(k) = 1$ for Negative Binomial distribution is not required.

Properties of Neg Binom $p(k) = \binom{k-1}{r-1}(1-p)^{k-r}p^r$

- Mean
$$E[X] = \frac{r}{p}$$
- Variance
$$Var[X] = \frac{r(1-p)}{p^2}$$
- Recall that for Geometric(p),
$$\mu = \frac{1}{p}, \sigma^2 = \frac{1-p}{p^2}$$

Pmf of Negative Binomial distribution



Question

Consider independent trials with success probability p. Let $q = 1 - p$. What's the probability of getting r successes before m failures?

(a) $p^{r-1}q^m$

(b) $\binom{r+m-1}{r-1}p^{r-1}q^m$

(c) $\sum_{k=r}^{r+m}\binom{k-1}{r-1}p^r q^{k-r}$

(d) $\sum_{k=r}^{r+m-1}\binom{k-1}{r-1}p^r q^{k-r}$

Let X denote the number of trials needed to get r successes. Then $X \sim \text{NB}(r, p)$.

$$
\begin{aligned}
&P(r \text{ successes before } m \text{ failures}) \\
=&P(r^{\text{th}} \text{ success occurs on trials } r, r+1, \ldots, r+m-1) \\
=&P(X = r) + P(X = r+1) + \cdots + P(X = r+m-1)
\end{aligned}
$$