STAT 24400 Lecture 3
Discrete Random Variables (Section 2.1)

Yibi Huang
Department of Statistics
University of Chicago

Random Variables

Random Variables

- So far we have considered probabilities for **events** (subsets) in a space space.
- But sample spaces are often "complicated", e.g.,
  - Coin tossing: a string of outcomes such as `TTHHTTTHTHTTTTH...`
  - Collecting responses for a survey: a long list of the answers to all the items: (Yes;1980;3;2000$;Chicago;No;1;Maybe;N/A;7;. . .)
- In most cases, we are interested in some specific numerical properties computed from the "outcome" itself, e.g.,
  - # of tosses required to get the first heads
  - # of people answered yes to item #5 in a survey.
- Such a numerical outcome from a random phenomenon is a *random variable*.

Random Variable

Formally speaking, a *random variable* is a real-valued function on the sample space Ω and maps elements of Ω, ω, to real numbers.

$$\begin{array}{ccc} \Omega & \xrightarrow{X} & \mathbb{R} \\ \omega & \longmapsto & x = X(\omega) \end{array}$$

Ex 1. Let X be the number of heads in 3 tosses of a coin. Sample space $\Omega = \{\mathtt{HHH}, \mathtt{HHT}, \mathtt{HTH}, \mathtt{HTT}, \mathtt{THH}, \mathtt{THT}, \mathtt{TTH}, \mathtt{TTT}\}$. Then

$$\begin{array}{llll} X(\mathtt{HHH}) = 3, & X(\mathtt{HHT}) = 2, & X(\mathtt{HTH}) = 2, & X(\mathtt{HTT}) = 1, \\ X(\mathtt{THH}) = 2, & X(\mathtt{THT}) = 1, & X(\mathtt{TTH}) = 1, & X(\mathtt{TTT}) = 0 \end{array}$$

Ex 2. Let Y be the number of tosses required to get a head. $\Omega = \{\mathtt{H}, \mathtt{TH}, \mathtt{TTH}, \mathtt{TTTH}, \mathtt{TTTTH}, \ldots\}$ Then

$$Y(\mathtt{H}) = 1,\ Y(\mathtt{TH}) = 2,\ Y(\mathtt{TTH}) = 3,\ Y(\mathtt{TTTH}) = 4, \ldots$$

Discrete and Continuous Random Variable

There are two types of random variables:

- *Discrete random variables* can only take a finite or countable infinite number of different values
  - Example: Number of heads obtained, number of batteries replaced last year
- *Continuous random variables* take real (decimal) values
  - Example: lifetime of a battery, someone's blood pressure

Distribution of a Discrete Random Variable

Coin Example

Let X = number of heads in 4 tosses of a fair coin.

$$\mathrm{P}(X = 0) = \mathrm{P}(\{\texttt{TTTT}\}) = 1/16$$

Coin Example

Let X = number of heads in 4 tosses of a fair coin.

$$\mathrm{P}(X = 0) = \mathrm{P}(\{\mathtt{TTTT}\}) = 1/16$$
$$\mathrm{P}(X = 1) = \mathrm{P}(\{\mathtt{HTTT}, \mathtt{THTT}, \mathtt{TTHT}, \mathtt{TTTH}\}) = 4/16$$

Coin Example

Let X = number of heads in 4 tosses of a fair coin.

$$\mathrm{P}(X = 0) = \mathrm{P}(\{\mathtt{TTTT}\}) = 1/16$$
$$\mathrm{P}(X = 1) = \mathrm{P}(\{\mathtt{HTTT}, \mathtt{THTT}, \mathtt{TTHT}, \mathtt{TTTH}\}) = 4/16$$
$$\mathrm{P}(X = 2) = \mathrm{P}(\{\mathtt{HHTT}, \mathtt{HTHT}, \mathtt{HTTH}, \mathtt{THHT}, \mathtt{THTH}, \mathtt{TTHH}\}) = 6/16$$

Coin Example

Let X = number of heads in 4 tosses of a fair coin.

$$\mathrm{P}(X = 0) = \mathrm{P}(\{\mathtt{TTTT}\}) = 1/16$$
$$\mathrm{P}(X = 1) = \mathrm{P}(\{\mathtt{HTTT}, \mathtt{THTT}, \mathtt{TTHT}, \mathtt{TTTH}\}) = 4/16$$
$$\mathrm{P}(X = 2) = \mathrm{P}(\{\mathtt{HHTT}, \mathtt{HTHT}, \mathtt{HTTH}, \mathtt{THHT}, \mathtt{THTH}, \mathtt{TTHH}\}) = 6/16$$
$$\mathrm{P}(X = 3) = \mathrm{P}(\{\mathtt{HHHT}, \mathtt{HHTH}, \mathtt{HTHH}, \mathtt{THHH}\}) = 4/16$$
$$\mathrm{P}(X = 4) = \mathrm{P}(\{\mathtt{HHHH}\}) = 1/16$$

Coin Example

Let X = number of heads in 4 tosses of a fair coin.

$$
\begin{aligned}
\mathrm{P}(X=0) &= \mathrm{P}(\{\mathtt{TTTT}\}) = 1/16 \\
\mathrm{P}(X=1) &= \mathrm{P}(\{\mathtt{HTTT}, \mathtt{THTT}, \mathtt{TTHT}, \mathtt{TTTH}\}) = 4/16 \\
\mathrm{P}(X=2) &= \mathrm{P}(\{\mathtt{HHTT}, \mathtt{HTHT}, \mathtt{HTTH}, \mathtt{THHT}, \mathtt{THTH}, \mathtt{TTHH}\}) = 6/16 \\
\mathrm{P}(X=3) &= \mathrm{P}(\{\mathtt{HHHT}, \mathtt{HHTH}, \mathtt{HTHH}, \mathtt{THHH}\}) = 4/16 \\
\mathrm{P}(X=4) &= \mathrm{P}(\{\mathtt{HHHH}\}) = 1/16
\end{aligned}
$$

The probability for each possible value of X is

Possible Value x of X	0	1	2	3	4
Probability $\mathrm{P}(X = x)$	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$

Note: these probabilities add up to 1:

$$\frac{1}{16} + \frac{4}{16} + \frac{6}{16} + \frac{4}{16} + \frac{1}{16} = 1$$

Probability Mass Function (PMF)

Probability Mass Function (PMF)

The *probability mass function* (PMF) of a random variable X is a function $p(x)$ that maps each possible value x_i to the corresponding probability $\mathrm{P}(X = x_i)$.

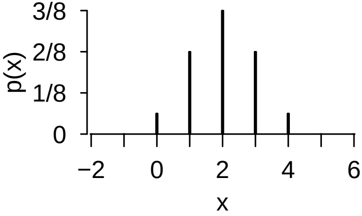
- A PMF $p(x)$ must satisfy $0 \leq p(x) \leq 1$ and $\sum_x p(x) = 1$.

Example (coin tossing on the previous slide)

Possible Values of X	0	1	2	3	4
Probabilities	$1/16$	$4/16$	$6/16$	$4/16$	$1/16$

The PMF of X is

$$p(x) = \begin{cases} 1/16 & \text{if } x = 0 \text{ or } 4 \\ 4/16 & \text{if } x = 1 \text{ or } 3 \\ 6/16 & \text{if } x = 2 \\ 0 & \text{if } x \neq 0, 1, 2, 3, 4 \end{cases}$$



Example: A Card Game

Consider a card game that you draw ONE card from a well-shuffled deck of cards. You win

- \$1 if you draw a heart,
- \$5 if you draw an ace (including the ace of hearts),
- \$10 if you draw the king of spades and
- \$0 for any other card you draw.

What's the PMF of your reward X?

Example: A Card Game

Consider a card game that you draw ONE card from a well-shuffled deck of cards. You win

- \$1 if you draw a heart,
- \$5 if you draw an ace (including the ace of hearts),
- \$10 if you draw the king of spades and
- \$0 for any other card you draw.

What's the PMF of your reward X?

Outcome	x	$p(x)$
Heart (not ace)	1	12/52
Ace	5	4/52
King of spades	10	1/52
All else	0	35/52

$$\Rightarrow \quad p(x) = \begin{cases} 35/52 & \text{if } x = 0 \\ 12/52 & \text{if } x = 1 \\ 4/52 & \text{if } x = 5 \\ 1/52 & \text{if } x = 10 \\ 0 & \text{for all other values of } x \end{cases}$$

Common Discrete Distributions

Bernouli Distribution

A random variable X that can take only two values, 0 and 1, with probabilities $1 - p$ and p, respectively, is called a *Bernoulli random variable*. Its PMF is thus

$$\begin{aligned} p(1) &= p \\ p(0) &= 1 - p \\ p(x) &= 0, \quad \text{if } x \neq 0 \text{ or } 1 \end{aligned}$$

Such a distribution is called *Bernoulli distribution* with parameter p.

Bernoulli Trials

A random trial having only 2 possible outcomes (S = Success, F = Failure) is called a *Bernoulli trial*, e.g.,

- whether a coin lands heads or tails when tossing a coin
- whether one gets a six or not a six when rolling a die
- whether a drug works on a patient or not
- whether a electronic device is defected
- whether a subject answers Yes or No to a survey question

If the probability of Success for a Bernoulli trial is $\mathrm{P}(S) = p$, and let $X = 1$ if the outcome is a Success and 0 if a Failure, then X would be a Bernoulli random variable with parameter p.

Binomial Distributions

Suppose n **independent** Bernoulli trials are to be performed, each of which results in

- a *success* with probability p and
- a *failure* with probability $1 - p$.

Define

$$X = \text{the number of successes obtained in the } n \text{ trials},$$

then X is said to have a *binomial distribution* with parameters (n, p), denoted as

$$X \sim Bin(n, p).$$

with the probability mass function (PMF)

$$\mathrm{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \ldots, n.$$

How the Binomial PMF above is obtained? (Next slide)

Consider the case with $n = 5$ trails. Possible outcomes for the event $X = 2$ (2 successes in 5 trials) are the $\binom{5}{2} = 10$ possible orderings of the 2 successes and 3 failures:

Possible Orders
$S\,S\,F\,F\,F$
$S\,F\,S\,F\,F$
$S\,F\,F\,S\,F$
$S\,F\,F\,F\,S$
$F\,S\,S\,F\,F$
$F\,S\,F\,S\,F$
$F\,S\,F\,F\,S$
$F\,F\,S\,S\,F$
$F\,F\,S\,F\,S$
$F\,F\,F\,S\,S$

As the trials are independent, the probabilities for the outcomes are respectively,

$$\begin{aligned}\mathrm{P}(SSFFF) &= \mathrm{P}(S)\mathrm{P}(S)\mathrm{P}(F)\mathrm{P}(F)\mathrm{P}(F)\\ &= pp(1-p)(1-p)(1-p) = p^2(1-p)^3,\\ \mathrm{P}(SFSFF) &= \mathrm{P}(S)\mathrm{P}(F)\mathrm{P}(S)\mathrm{P}(F)\mathrm{P}(F)\\ &= p(1-p)p(1-p)(1-p) = p^2(1-p)^3,\\ &\text{etc}\end{aligned}$$

Observe the 10 outcomes have equal probability $p^2(1-p)^3$ since they all have 2 Successes and 3 Failures.

As the outcomes above are disjoint, $\mathrm{P}(X = 2) = \mathrm{P}(2$ successes in 5 trials) is the sum of their probabilities

$$\mathrm{P}(X = 2) = \binom{5}{2} p^2(1-p)^3.$$

In general, for $X \sim \text{Bin}(n, p)$, outcomes in the event $\{X = k\} = \{k \text{ successes in } n \text{ trials}\}$ are the $\binom{n}{k}$ possible orderings of the k successes and $n - k$ failures that each has probability $p^k(1-p)^{n-k}$ to occur. The Binomial PMF is thus

$$p(k) = \mathrm{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \ldots, n.$$

Does the Binomial PMF Add Up to 1?

A legitimate PMF $p(k)$ must add up to 1 $\sum_k p(k) = 1$.
Does the Binomial PMF satisfy the condition

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

Does the Binomial PMF Add Up to 1?

A legitimate PMF $p(k)$ must add up to 1 $\sum_k p(k) = 1$.
Does the Binomial PMF satisfy the condition

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

Yes, using the Binomial Expansion

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k},$$

plugging in $a = p$ and $b = 1 - p$, we get

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

Sum of i.i.d. Bernoulli Random Variables is Binomial

If $X_1, X_2, \ldots, X_n$ are i.i.d. Bernoulli random variables with success probability p, then

$$X_1 + X_2 + \ldots + X_n \sim \text{Bin}(n, p).$$

where "i.i.d." $=$ independent and identically distributed.

Sum of i.i.d. Bernoulli Random Variables is Binomial

If $X_1, X_2, \ldots, X_n$ are i.i.d. Bernoulli random variables with success probability p, then

$$X_1 + X_2 + \ldots + X_n \sim \text{Bin}(n, p).$$

where "i.i.d." $=$ independent and identically distributed.

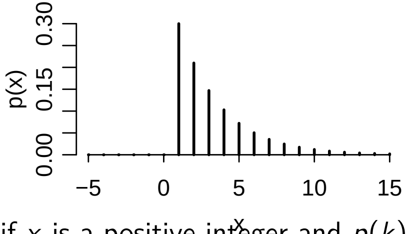
Moreover, if $X \sim \text{Bin}(m, p)$ and $Y \sim \text{Bin}(\ell, p)$ are independent, then

$$X + Y \sim \text{Bin}(m + \ell, p).$$

Geometric Distribution

Suppose that a sequence of independent Bernoulli trials are performed, each with probability of success p. Let X be the number of trials required to obtain the first Success.

The PMF of X is



$$\begin{aligned} p(k) = \mathrm{P}(X = k) &= \mathrm{P}(\overbrace{F \dots F}^{k-1 \text{ F's}} S) \quad \text{by indep.} \\ &= \mathrm{P}(F) \cdots \mathrm{P}(F)\mathrm{P}(S) \\ &= \underbrace{(1-p)\cdots(1-p)}_{k-1 \text{ copies}} p \\ &= (1-p)^{k-1} p, \end{aligned}$$

if x is a positive integer and $p(k) = 0$ if not, denoted as

$$X \sim \text{Geometric}(p).$$

We say X has a *geometric distribution*, since the PMF is a geometric sequence.

Does the Geometric PMF Add Up to 1?

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

Does the Geometric PMF Add Up to 1?

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

Recall the geometric series

$$\begin{aligned} \sum_{k=0}^{\infty} ax^k &= a + ax + ax^2 + \cdots ax^k + \cdots \\ &= \frac{a}{1-x} \quad \text{if } |x| < 1. \end{aligned}$$

Does the Geometric PMF Add Up to 1?

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

Recall the geometric series

$$\begin{aligned}\sum_{k=0}^{\infty} ax^k &= a + ax + ax^2 + \cdots ax^k + \cdots \\ &= \frac{a}{1-x} \quad \text{if } |x| < 1.\end{aligned}$$

The sum of the Geometric PMF

$$\begin{aligned}\sum_{k=1}^{\infty} p(k) &= \sum_{k=1}^{\infty}(1-p)^{k-1}p \\ &= p + (1-p)p + (1-p)^2 p + \cdots + (1-p)^{k-1}p + \cdots\end{aligned}$$

is simply the case that $a = p$ and $x = 1 - p$ and hence the sum is

$$\frac{a}{1-r} = \frac{p}{1-(1-p)} = \frac{p}{p} = 1.$$

Negative Binomial Distributions

Suppose that a sequence of independent Bernoulli trials are performed, each with probability of success p. Let X be the number of trials required to obtain the rth Success.

For the event $\{X = k\}$ to occur,

- the kth trial must be a Success,
- the first $k-1$ trials can be $r-1$ successes and $k-r$ failures in any order.

$$\underbrace{\overbrace{\underline{\ ?\ }\ \underline{\ ?\ }\ \cdots\cdots\ \underline{\ ?\ }}^{\text{first } k-1 \text{ trials}}}_{\substack{r-1 \text{ S's and } k-r \text{ F's} \\ \text{in any order}}}\quad \underbrace{\overbrace{\underline{\ S\ }}^{k\text{th trial}}}_{r\text{th S}}$$

Thus, the Negative Binomial PMF is

$$\begin{aligned} \mathrm{P}(X = k) &= \overbrace{\binom{k-1}{r-1} p^{r-1}(1-p)^{k-r}}^{\text{first } k-1 \text{ trials}} \times \overbrace{p}^{k\text{th trial}} \\ &= \binom{k-1}{r-1} p^{r}(1-p)^{k-r} \quad k = r, r+1, \ldots . \end{aligned}$$

denoted as $X \sim NB(r, p)$.

Relation Between Negative Binomial & Geometric

If $X_1, X_2, \ldots, X_r$ are i.i.d. random variables with a Geometric(p) distribution, then

$$X_1 + X_2 + \ldots + X_r \sim \mathrm{NB}(r, p).$$

Relation Between Negative Binomial & Geometric

If $X_1, X_2, \ldots, X_r$ are i.i.d. random variables with a Geometric(p) distribution, then

$$X_1 + X_2 + \ldots + X_r \sim \text{NB}(r, p).$$

Conversely, let

- X_1 be the number of trials needed to get the first Success
- X_2 be the number of additional trials needed to get the 2nd Success after the first Success
- $\vdots$
- X_r be the number of additional trials needed to get the rth Success after the $(r-1)$st Success

then $X_1, X_2, \ldots, X_r$ are independent Geometric(p) random variables.

Negative Binomial Expansion (1)

Recall the geometric series: for $|x| < 1$,

$$\frac{1}{1-x} = 1 + x + x^2 + \cdots x^k + \cdots = \sum_{k=0}^{\infty} x^k.$$

Taking derivative on both sides, we get

$$\frac{1}{(1-x)^2} = 1 + 2x + \cdots kx^{k-1} + \cdots = \sum_{k=1}^{\infty} \binom{k}{1} x^{k-1}.$$

Taking the 2nd derivative on both sides, we get

$$\frac{2}{(1-x)^3} = \sum_{k=2}^{\infty} \binom{k}{1}(k-1)x^{k-2} = \sum_{k=2}^{\infty} 2\binom{k}{2} x^{k-2}$$

Dividing both sides by 2, we get

$$\frac{1}{(1-x)^3} = \sum_{k=2}^{\infty} \binom{k}{2} x^{k-2}.$$

Negative Binomial Expansion (2)

By Mathematical Induction, one can show that the mth derivative of $\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k$ is

$$\frac{1}{(1-x)^{m+1}} = \sum_{k=m}^{\infty} \binom{k}{m} x^{k-m},$$

called the *Negative Binomial expansion*.

The sum of the Negative Binomial PMF can be obtained by apply the Negative Binomial expansion with $m = r - 1$ and $x = 1 - p$,

$$\begin{aligned}\sum_{k=r}^{\infty} \binom{k-1}{r-1} p^r (1-p)^{k-r} &= p^r \sum_{k=r}^{\infty} \binom{k-1}{r-1} (1-p)^{k-r} \\ &= p^r \cdot \frac{1}{(1-(1-p))^{r-1+1}} = \frac{p^r}{p^r} = 1.\end{aligned}$$

Hypergeometric Distribution

R red balls, B blue balls

↙ ↘

d balls drawn at random without replacement: X red, $d-X$ blue

$R+B-d$ balls remained in the original box: $R-X$ red, $B-(d-X)$ blue

Suppose d draws are made at random w/o replacement from a box containing R red balls and B blue balls. The number of red balls X obtained in d draws has a **hypergeometric distribution:**

$$\mathrm{P}(X=x) = \mathrm{P}(x \text{ red},\ d-x \text{ blue})$$

$$= \frac{\binom{\#\text{ of ways to pick } x \text{ red balls}}{\text{out of R red balls}}\binom{\#\text{ of ways to pick } d-x \text{ blue balls}}{\text{out of B blue balls}}}{(\#\text{ of ways to pick } d \text{ balls out of } R+B \text{ balls})} = \frac{\binom{R}{x}\binom{B}{d-x}}{\binom{R+B}{d}}$$

for $0 \le x \le R$, $0 \le d-x \le B$.

Hypergeometric Distribution

$$\boxed{R \text{ red balls}, \quad B \text{ blue balls}}$$

$\swarrow$ $\searrow$

d balls drawn at random without replacement	$R+B-d$ balls remained in the original box
X red, $d-X$ blue	$R-X$ red, $B-(d-X)$ blue

Suppose $\boxed{d \text{ draws}}$ are made at random w/o replacement from a box containing R red balls and B blue balls. The number of red balls X obtained in d draws has a **hypergeometric distribution:**

$$\mathrm{P}(X=x)=\mathrm{P}(x \text{ red}, \ d-x \text{ blue})$$

$$=\frac{\binom{\#\text{ of ways to pick } x \text{ red balls}}{\text{out of R red balls}}\binom{\#\text{ of ways to pick } d-x \text{ blue balls}}{\text{out of B blue balls}}}{(\#\text{ of ways to pick } d \text{ balls out of } R+B \text{ balls})}=\frac{\binom{R}{x}\binom{B}{d-x}}{\binom{R+B}{d}}$$

for $0 \le x \le R$, $0 \le d-x \le B$.

Q: If the draws are made with replacement, what's the distribution of X?

Hypergeometric Distribution

$$\boxed{R \text{ red balls}, \quad B \text{ blue balls}}$$

$\swarrow$ $\searrow$

d balls drawn at random without replacement	$R+B-d$ balls remained in the original box
X red, $d-X$ blue	$R-X$ red, $B-(d-X)$ blue

Suppose $\boxed{d \text{ draws}}$ are made at random w/o replacement from a box containing R red balls and B blue balls. The number of red balls X obtained in d draws has a **hypergeometric distribution:**

$$\mathrm{P}(X=x) = \mathrm{P}(x \text{ red}, \ d-x \text{ blue})$$

$$= \frac{\binom{\#\text{ of ways to pick } x \text{ red balls}}{\text{out of R red balls}}\binom{\#\text{ of ways to pick } d-x \text{ blue balls}}{\text{out of B blue balls}}}{(\#\text{ of ways to pick } d \text{ balls out of } R+B \text{ balls})} = \frac{\binom{R}{x}\binom{B}{d-x}}{\binom{R+B}{d}}$$

for $0 \le x \le R$, $0 \le d-x \le B$.

Q: If the draws are made with replacement, what's the distribution of X? $X \sim \text{Bin}\left(d, \, p = \frac{R}{R+B}\right)$.

The hypergeometric PMF adds up to 1

$$\sum_x \mathrm{P}(X = x) = \frac{\sum_x \binom{R}{x}\binom{B}{d-x}}{\binom{R+B}{d}} = 1$$

because of **Vandermonde's identity** of the Binomial coefficients.

$$\binom{m+n}{r} = \sum_{k=0}^{r} \binom{m}{k}\binom{n}{r-k}.$$

Hypergeometric $\approx$ Binomial if ...

If the number of balls are large (R and B both $\to \infty$ and $\frac{R}{R+B} \to p$), then

$$\mathrm{P}(X = x) = \frac{\binom{R}{x}\binom{B}{d-x}}{\binom{R+B}{d}} \longrightarrow \binom{d}{x} p^x (1-p)^{d-x} \quad \text{for } x = 0, 1, \ldots, d,$$

i.e., drawing with or without replacement makes little difference.

Proof.

$$\frac{\binom{R}{x}\binom{B}{d-x}}{\binom{R+B}{d}} = \frac{\frac{R!}{x!(R-x)!}\frac{B!}{(d-x)!(B-d+x)!}}{\frac{(R+B)!}{d!(R+B-d)!}} = \frac{d!}{x!\,(d-x)!}\frac{\frac{R!}{(R-x)!}\frac{B!}{(B-d+x)!}}{\frac{(R+B)!}{(R+B-d)!}}$$

$$= \binom{d}{x} \underbrace{\frac{R}{R+B}}_{\to p} \underbrace{\frac{R-1}{R+B-1}}_{\to p} \cdots \underbrace{\frac{R-x+1}{R+B-x+1}}_{\to p}$$

$$\times \underbrace{\frac{B}{R+B-x}}_{\to 1-p} \underbrace{\frac{B-1}{R+B-x-1}}_{\to 1-p} \cdots \underbrace{\frac{B-d+x+1}{(R+B-d+1)}}_{\to 1-p}$$

$$\to \binom{d}{x} p^x (1-p)^{d-x}.$$

Poisson Distribution

A random variable X has a Poisson distribution with parameter $\lambda > 0$ if its PMF is

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

denoted as

$$X \sim \text{Poisson}(\lambda).$$

We can show Poisson PMF sum to 1 using the Taylor expansion of the exponential function: $e^u = \sum_{k=0}^{\infty} u^k / k!$, and obtain

$$\sum_{k=0}^{\infty} \mathrm{P}(X = k) = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \underbrace{\sum_{k=0}^{\infty} \frac{\lambda^k}{k!}}_{=e^{\lambda}} = e^{-\lambda} e^{\lambda} = 1.$$

Poisson Approximation to Binomial

For a **Binomial** distribution with huge n and tiny p such that np moderate,

$$\text{Binomial}(n, p) \text{ is approx. Poisson}(\lambda = np).$$

Below are the values of $\mathrm{P}(Y = k)$, $k = 0, 1, 2, 3, 4, 5$ for

$$Y \sim \text{Binomial}(n = 50,\ p = 0.03),\ \text{and}$$
$$Y \sim \text{Poisson}(\lambda = 50 \times 0.03 = 1.5).$$

```
dbinom(0:5, size=50, p=0.03)         # Binomial(n=50, p=0.03)
[1] 0.21807 0.33721 0.25552 0.12644 0.04595 0.01307
dpois(0:5, lambda = 50*0.03)         # Poisson(lambda = 50*0.03)
[1] 0.22313 0.33470 0.25102 0.12551 0.04707 0.01412
```

Proof of Poisson Approximation to Binomial

The Binomial PMF is

$$
\begin{aligned}
\mathrm{P}(X = k) &= \binom{n}{k} p^k (1-p)^{n-k} \\
&= \frac{n!}{k!(n-k)!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \quad \text{(setting } \lambda = np\text{)} \\
&= \frac{\lambda^k}{k!} \underbrace{\frac{n!}{(n-k)!n^k}}_{\to 1} \underbrace{\left(1 - \frac{\lambda}{n}\right)^n}_{\to e^{-\lambda}} \underbrace{\left(1 - \frac{\lambda}{n}\right)^{-k}}_{\to 1} \\
&\to \frac{\lambda^k}{k!} e^{-\lambda} \quad \text{as } n \to \infty \text{ with } np \to \lambda.
\end{aligned}
$$

Example — Fatalities From Horse Kicks (p.45, Textbook)

The # of deaths in a year resulted from being kicked by a horse or mule was recorded for each of 10 corps of Prussian cavalry over a period of 20 years, giving 200 corps-years worth of data.

# of Deaths (in a corp in a year)	0	1	2	3	4	Total
Frequency	109	65	22	3	1	200

The count of deaths due to horse kicks in a corp in a given year may have a Poisson distribution because

- $p = P(\text{a soldier died from horsekicks in a given year}) \approx 0$;
- $n = \#$ of soldiers in a corp was large (100's or 1000's);
- whether a soldier was kicked was (at least nearly) independent of whether others were kicked

Example (Fatalities From Horse Kicks — Cont'd)

The fitted Poisson probability to have k deaths from horsekicks for $\lambda = 0.61$ is

$$\mathrm{P}(Y = k) = e^{-\lambda}\frac{\lambda^k}{k!} = e^{-0.61}\frac{(0.61)^k}{k!}, \quad k = 0, 1, 2, \ldots.$$

k	Observed Frequency	Relative Frequency	Poisson Probability
0	109	0.545	0.543
1	65	0.325	0.331
2	22	0.110	0.101
3	3	0.015	0.021
4	1	0.005	0.003
Total	200	1	0.999

- $\lambda = 0.61$ is the average of the 200 counts

$$\frac{0 \times 109 + 1 \times 65 + 2 \times 22 + 3 \times 3 + 4 \times 1}{200} = 0.61.$$

When Do Poisson Distributions Come Up?

Variables that are generally Poisson:

- # of misprints on a page of a book
- # of calls coming into an exchange during a unit of time (if the exchange services a large number of customers who act more or less independently.)
- # of people in a community who survive to age 100
- # of vehicles that pass a marker on a roadway during a unit of time (for light traffic only. In heavy traffic, however, one vehicle's movement may influence another)

Exercise 1

If you roll a fair die 50 times, what is the distribution of the # of ⚀'s rolled? And, how likely is it that you will get no more than 5 ⚀'s?

Exercise 1

If you roll a fair die 50 times, what is the distribution of the # of ⚀'s rolled? And, how likely is it that you will get no more than 5 ⚀'s?

Let X = number of ⚀'s. Then $X \sim \text{Binomial}(n = 50,\ p = 1/6)$.

$$\mathrm{P}(X \leq 5) = \sum_{k=0}^{5} \mathrm{P}(X = k) = \sum_{k=0}^{5} \binom{50}{k} \left(\frac{1}{6}\right)^{k} \left(\frac{5}{6}\right)^{50-k} \approx 0.139.$$

Exercise 2

Suppose you draw 10 cards from a standard deck without replacement.
Let X be the # of Kings you draw.
What's the distribution of X?

Exercise 2

Suppose you draw 10 cards from a standard deck without replacement.
Let X be the $\#$ of Kings you draw.
What's the distribution of X?

X has a hypergeometric distribution where

- the Red balls are the 4 Kings
- the Blue balls are the remaining 48 cards that are not Kings

The PMF is

$$\mathrm{P}(X = k) = \frac{\binom{4}{k}\binom{48}{10-k}}{\binom{52}{10}}, \quad \text{for } k = 0, 1, 2, 3, 4.$$

Exercise 3

Consider the following game:
At each round, you roll a red die & a blue die.
If the red die is even, you win a prize, otherwise you win nothing.
If the blue die is a 1, then you stop playing, otherwise you continue.

1. What is the distribution of $X =$ the $\#$ of rounds you play?
2. What is the distribution of the $Y =$ of times you win?

Ans.

1. $X = \#$ of rolls needed to reach the first 1, so
   $X \sim \text{Geometric}(1/6)$.

$$\mathrm{P}(X = x) = (5/6)^{x-1}(1/6).$$

Exercise 3 (Cont'd)

Recall $X = \#$ rounds you play, and $Y = \#$ times you win.

Observe that given $X = x$, $Y \sim \text{Bin}(x, 1/2)$.

To calculate the PMF of Y, we can use the Law of Total Probability, for each $y = 0, 1, 2, \ldots$,

$$\begin{aligned}
\mathrm{P}(Y = y) &= \sum_{x=1}^{\infty} \mathrm{P}(Y = y \text{ and } X = x) \\
&= \sum_{x=1}^{\infty} \mathrm{P}(X = x) \cdot \mathrm{P}(Y = y \mid X = x) \\
&= \sum_{x=1}^{\infty} \underbrace{\left(\tfrac{5}{6}\right)^{x-1} \cdot \tfrac{1}{6}}_{\text{from Geom. distrib.}} \cdot \underbrace{\binom{x}{y} \left(\tfrac{1}{2}\right)^{y} \left(\tfrac{1}{2}\right)^{x-y}}_{\text{from Binomial distrib.}} \\
&= \sum_{x=\max(1,y)}^{\infty} \left(\tfrac{5}{6}\right)^{x-1} \left(\tfrac{1}{6}\right) \binom{x}{y} \left(\tfrac{1}{2}\right)^{x}
\end{aligned}$$

For $y = 0$

$$\mathrm{P}(Y = 0) = \sum_{x=1}^{\infty} (\frac{5}{6})^{x-1} (\frac{1}{6}) (\frac{1}{2})^{x} = \frac{1}{12} \sum_{x=1}^{\infty} (\frac{5}{12})^{x-1} = \frac{1}{12} \frac{1}{(1 - 5/12)} = \frac{1}{7}.$$

For $y = 1, 2, 3, \ldots$,

$$\mathrm{P}(Y = y) = \sum_{x=y}^{\infty} (\frac{5}{6})^{x-1} (\frac{1}{6}) \binom{x}{y} (\frac{1}{2})^{x} = \frac{1}{5} (\frac{5}{12})^{y} \sum_{x=y}^{\infty} \binom{x}{y} (\frac{5}{12})^{x-y}$$

$$= \frac{1}{5} (\frac{5}{12})^{y} \frac{1}{(1 - 5/12)^{y+1}} = \frac{12}{35} (\frac{5}{7})^{y}.$$