

Week 4. Discrete Distributions: *pmf* and Expected Value

Idea:

A **random variable** X is a set of values x , with a probability $f(x)$ for each. Usually, discrete random variables **COUNT** things.

Probability Properties.

- P is a number
(fraction, decimal, percent)
- $0 \leq P \leq 1$
(for percent, $P \leq 100\%$)
- Total probability is 1
(for percent total is 100%)



pmf Properties.

- $0 \leq f(x) \leq 1$ for all x
- Sum of all values is 1
 $f(x_1) + f(x_2) + \dots = 1$

Table of **counts**.

x	1	3	4	Total
#	3	5	2	10

Data: 1, 1, 1, 3, 3, 3, 3, 3, 4, 4
 $1 (\times 3), 3 (\times 5), 4 (\times 2)$

Mean: $\frac{1 \times 3 + 3 \times 5 + 4 \times 2}{10}$



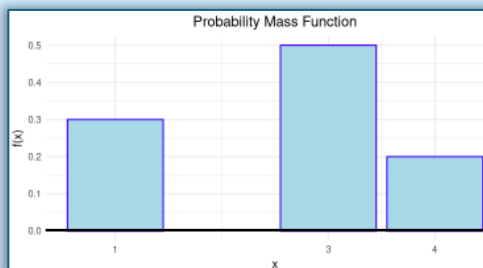
Table of **probabilities** (“proportions” or “frequencies”).

x	1	3	4
$P(X=x)$	$\frac{3}{10}$	$\frac{5}{10}$	$\frac{2}{10}$

Mean: $1 \times \frac{3}{10} + 3 \times \frac{5}{10} + 4 \times \frac{2}{10}$



Probability Mass Function (*pmf*)



$$f(1) = \frac{3}{10}, f(3) = \frac{5}{10}, f(4) = \frac{2}{10}$$

Expected Value: (“mean” or “ μ ”)

$$E[X] = 1 \times \frac{3}{10} + 3 \times \frac{5}{10} + 4 \times \frac{2}{10}$$

For each outcome x ,
pmf has $f(x) = P(X=x)$

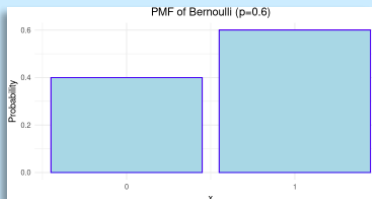
Expected Value is

$$E[X] = x_1 \cdot f(x_1) + x_2 \cdot f(x_2) + \dots$$

Week 4. Basic Discrete Distributions

Common patterns of pmf are organized into families called “distributions”

Bernoulli p



Bernoulli “trial”

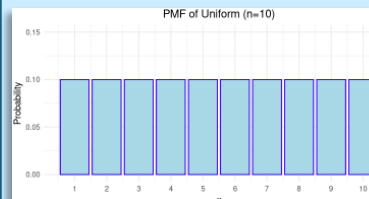
Check a binary event.

$$X = \begin{cases} 1 & \text{if success} \\ 0 & \text{if failure} \end{cases}$$

p = Probability of “success”

$1 - p$ = Probability of “failure”

Uniform n

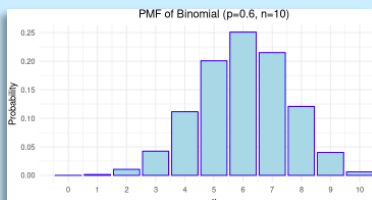


Pick randomly from n possible choices, each of them equally likely

$$f(x_i) = \frac{1}{n} \text{ for each } x_i$$

Expected value: $E[X] = \text{average } x_i$

Binomial n, p



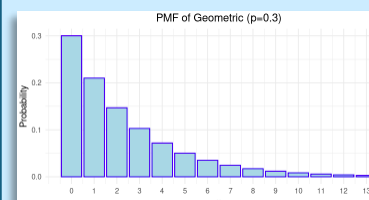
Sum of n indep. Bernoulli trials

Count **#success** in n attempts

p = Probability of “success”

Expected value: $E[X] = n \cdot p$

Geometric p



Waiting time with Bernoulli trials

Count **#trials** until success

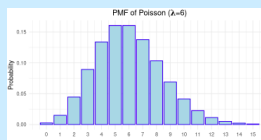
p = Probability of “success”

Expected value: $E[X] = 1/p$

“Memoryless” distribution

If a random variable follows one of these patterns, we say “ X is distributed as...”

Poisson λ

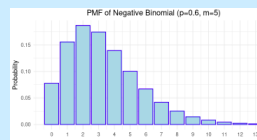


Similar to
binomial

Count **#success** in time period
 λ = rate / density of “success”

Expected value: $E[X] = \lambda$

Negative Binomial m, p



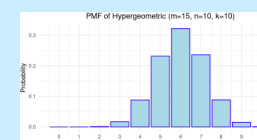
Similar to
geometric

Count **#failures** until m successes

p = Probability of “success”

Expected value: $E[X] = m/p - m$

Hypergeometric n, M, N



Binomial but
Bernoulli trials
are **not** indep.

Count **#success** in n “draws”
without replacement
from collection with N successes
and M failures

Used for small populations