6.1 Discrete Random Variables

pages 347–350

What is a random variable? → x_1

Takes numerical values that describe the outcome of some chance process.

What is a probability distribution? → p_1

It's a model for a random variable. Gives the possible values and their probability of occuring.

What is a discrete random variable?

- Random variable with a fixed set of probabilities with gaps in between.

  ex, dice rolls, coin flips, AP Score, shoe size

Imagine selecting a U.S. high school student at random. Define the random variable X = number of languages spoken by the randomly selected student. The table below gives the probability distribution of X, based on a sample of students from the U.S. Census at School database.

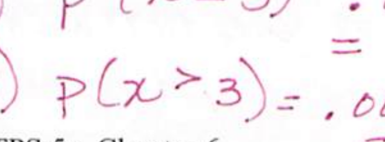
random variable → value → probability →

Languages:	1	2	3	4	5
Probability:	0.630	0.295	0.065	0.008	0.002

(a) Show that the probability distribution for X is legitimate.
(b) Make a histogram of the probability distribution. Describe what you see.
(c) What is the probability that a randomly selected student speaks at least 3 languages? Interpret this probability.
(d) What is the probability that a randomly selected student speaks more than 3 languages? How is this different than (c)?
(e) Given that a student speaks more than one language, what is the probability the student speaks 3 languages?

To be legit:

a) ① All probabilities must add up to 1. ✓
② Each probability has a value between zero & 1 ✓

b) 

V → 1 to 5 languages
S → Right skew
C → ~1 (median)
O → None

c) $P(x \geq 3) = .065 + .008 + .002$
$= .075$

d) $P(x > 3) = .008 + .002$
$= .01$

e) $\frac{.065}{.370}$

One wager players can make in Roulette is called a "corner bet." To make this bet, a player places his chips on the intersection of four numbered squares on the Roulette table. If one of these numbers comes up on the wheel and the player bet \$1, the player gets his \$1 back plus \$8 more. Otherwise, the casino keeps the original \$1 bet. If X = net gain from a single \$1 corner bet, the possible outcomes are $x = -1$ or $x = 8$. Here is the probability distribution of X:

Value x	–\$1	\$8
Probability $p(x)$	34/38	4/38

If a player were to make this \$1 bet over and over, what would be the player's average gain?

If player played 38 times

$$\mu_x = \frac{-1+-1+-1+-1+\ldots\ldots+8+8+8+8}{38}$$

$$= \frac{34(-1)+4(8)}{38} = -1\left(\frac{34}{38}\right)+8\left(\frac{4}{38}\right)$$

$= \$-.05$

Interpretation:
If a player makes \$1 bets many, many times, the approximate loss would be ~\$.05 per bet.

pages 350–352

How do you calculate the mean (expected value) of a discrete random variable? Is the formula on the formula sheet?

YES

$E(X) = \mu_x = \sum x_i p_i$ * sum of value * probability

How do you interpret the mean (expected value) of a discrete random variable?

Must include:
① "many repetitions/trials"
② "average"/"approx"/"estimate" ≈

Calculate and interpret the mean number of languages in the example on the previous page.

$\mu_x = 1(.63)+2(.295)+3(.065)+4(.008)+5(.002)$
$= 1.457$

If we randomly select many, many U.S. high school students, the average number of languages spoken would be approximately 1.457

Does the expected value of a random variable have to equal one of the possible values of the random variable? Should expected values be rounded?

No. E(x) is a long run probability of the average (mean) expected outcome.

DO NOT round E(x) to a whole number.

6.1 Discrete and Continuous Random Variables

pages 352–354

Suppose that X is a discrete random variable with probability distribution to the right, and that μ_x is the mean of X.

Value:	x_1	x_2	x_3	...
Probability:	p_1	p_2	p_3	...

Variance of X:

$Var(x) = \sigma_x^2 = (x_1-\mu_x)^2 p_1 + (x_2-\mu_x)^2 p_2 + \ldots$

$\sigma_x^2 = \sum (x_i-\mu_x)^2 p_i$

Standard Deviation of X:

$\sigma_x = \sqrt{(x_1-\mu_x)^2 p_1 + (x_2-\mu_x)^2 p_2 + \ldots}$

$= \sqrt{\sum (x_i-\mu_x)^2 p_i}$

How do you interpret the standard deviation of a discrete random variable?

"On average, the (x-variable) will vary from the mean of ____ by about ____."

Calculate and interpret the standard deviation of X in the languages example.

$Var(x) = \sigma_x^2 = (1-1.457)^2(.630) + (2-1.457)^2(.295) + \ldots = 0.450$

$\sigma_x = \sqrt{.450} = 0.671$

On average, the number of languages a randomly selected student speaks will vary from the mean of 1.457 by about 0.671.

How to calculate **mean and standard deviation** of a discrete random variable on the **calculator**:

- Enter x-values into list 1
- Enter probability values into list 2
- Calculate the 1-variable stats with list 1 in the list category
- Use list 2 as the frequency list
- Calculate

$\bar{x} = mean$, $S_x =$ standard deviation

In 2010, there were 1319 games played in the National Hockey League's regular season. Imagine selecting one of these games at random and then randomly selecting one of the two teams that played in the game. Define the random variable X = number of goals scored by a randomly selected team in a randomly selected game. The table below gives the probability distribution of X:

L1	Goals:	0	1	2	3	4	5	6	7	8	9
L2	Probability:	0.061	0.154	0.228	0.229	0.173	0.094	0.041	0.015	0.004	0.001

Sample: Compute the mean and standard deviation of the random variable X and interpret this value in context.

Using 1-Var-Stats. $\bar{x} = 2.851$

$\sigma_x = 1.632$

→ After viewing many, many NHL games, we would expect the average goals to be about 2.851.

→ On average, the goals scored will vary from the mean of 2.851 by about 1.632.

In an experiment on the behavior of young children, each subject is placed in an area with five toys. Past experiments have shown that the probability distribution of the number X of toys played with by a randomly selected subject is as follows:

X toys	0	1	2	3	4	5
Probability	0.03	0.16	0.30	0.23	0.17	0.11

Sample: Calculate the mean and standard deviation of the random variable X and interpret this value in context.

$\bar{x}$ = mean = $E_x \rightarrow \approx 2.68$

σ_x = standard deviation → ≈ 1.31

→ After observing many, many young children, we would expect the average number of toys played with to be about 2.68.

→ On average, the number of toys played with will vary from the mean of 2.68 by about 1.31

pages 355–358

\# Probability Models:
Discrete Random Variable assigns prob btwn 0 & 1 to each X value
Continuous RV - assigns prob of 0 to each outcome.

What is a continuous random variable?

• Takes on any value in an interval of values.
• often a situation where something is measured.

Is it possible to have a shoe size = 8? Is it possible to have a foot length = 8 inches?

yes   8 = 8.0000 inches → so unlikely

How many possible foot lengths are there? How can we graph the distribution of foot length?

infinite possibilities
Histogram with really skinny bins that add to 1 → density curve.

How do we find probabilities for continuous random variables?

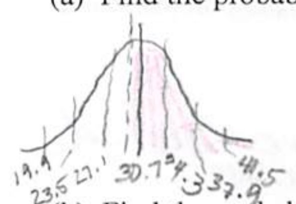
Area Under A Density Curve

For a continuous random variable X, how is $P(X < a)$ related to $P(X \le a)$?

$P(X<a)$, $P(X \le a)$ same for continuous because the boundary line adds no extra area

The weights of three-year-old females closely follow a Normal distribution with a mean of $\mu = 30.7$ pounds and a standard deviation of $\sigma = 3.6$ pounds. Randomly choose one three-year-old female and call her weight X.

(a) Find the probability that the randomly selected three-year-old female weighs at least 30 pounds.



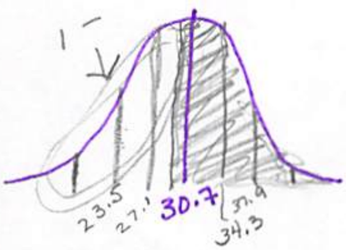
$P(x \ge 30) = P(x > 30)$
normal CDF ($LB = 30$, $UB = 100$, $\mu = 30.7$, $\sigma = 3.6$) $= 0.5771$
Probability that a randomly selected 3yr old female weighs at least 30 lbs is .5771

(b) Find the probability that a randomly selected three-year-old female weighs between 25 and 35 pounds.

(c) If $P(X < k) = 0.8$, find the value of k.

The weights of three-year-old females closely follow a Normal distribution with a mean of μ = 30.7 pounds and a standard deviation of σ = 3.6 pounds. Randomly choose one three-year-old female and call her weight X.

(a) Find the probability that the randomly selected three-year-old female weighs at least 30 pounds.



normalcdf with lower bound = 30, upperbound = 100, μ = 30.7, σ = 3.6

$P(x \geq 30) = p(x > 30)$

$= p\left(z > \frac{30-30.7}{3.6}\right) = P(z > -0.194)$

$1 - p(z < -.194)$

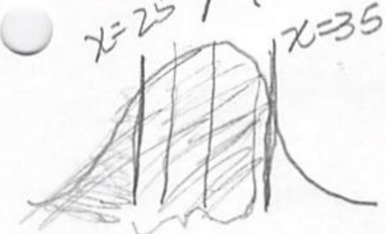
$1 - .4247$

$= 0.5753$

About 57.5% chance that a randomly selected 3 year old female weighs at least 30 lbs.

(b) Find the probability that a randomly selected three-year-old female weighs between 25 and 35 pounds.

$P(25 \leq x \leq 35)$



$P\left(z < \frac{35-30.7}{3.6}\right) - P\left(z < \frac{25-30.7}{3.6}\right)$

normalcdf(25, 35, 30.7, 3.6) (LB, UB, μ, σ)

$= 0.8272$

(c) If $P(X < k) = 0.8$, find the value of k.

Area closest to .8 → $z = 0.84$

$\frac{K - 30.7}{3.6} = .84$

$K = 33.724$

$K = \text{invNorm}(\text{area} = .8, \mu = 30.7, \sigma = 3.6)$

$K = 33.73$