Lecture 30: Probability Theory

Chapter 2.4-2.5

Random Variable

Intuitively Thinking: Expected Value

Intuitively Thinking: Expected Value

Say you have a random variable X:

x	2	3	4	10	11
$\Pr(X = x)$	$\frac{15}{100}$	$\frac{25}{100}$	$\frac{10}{100}$	$\frac{30}{100}$	$\frac{20}{100}$

E.g. We observe $X = 3$ with prob .25

Intuitively Thinking: Expected Value

Say you have a random variable X:

x	2	3	4	10	11
$\Pr(X = x)$	$\frac{15}{100}$	$\frac{25}{100}$	$\frac{10}{100}$	$\frac{30}{100}$	$\frac{20}{100}$

E.g. We observe $X = 3$ with prob .25

Is the value we expect to observe:

$$\frac{2+3+4+10+11}{5} \quad = \quad 6 \;?$$

Intuitively Thinking: Expected Value

No, each of the x's have different probability of occurring.

Intuitively Thinking: Expected Value

No, each of the x's have different probability of occurring.

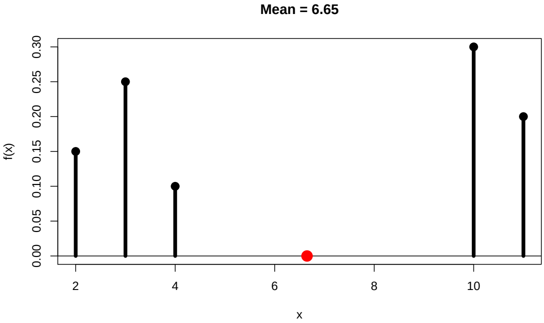
For each x, we assign different weights $\Pr(X = x)$ and not $\frac{1}{5}$:

$$2 \times \frac{15}{100} + 3 \times \frac{25}{100} + 4 \times \frac{10}{100} + 10 \times \frac{30}{100} + 11 \times \frac{20}{100} = 6.65$$

Expected Value

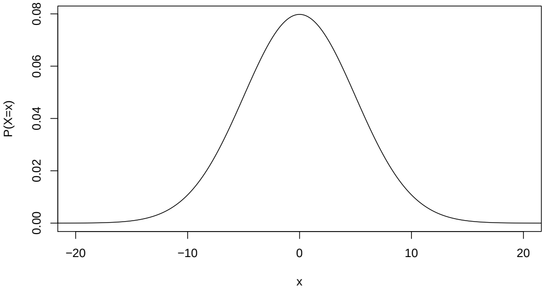
Expected Value

You can also think of the mean as the center of mass or balance point (marked with red point):



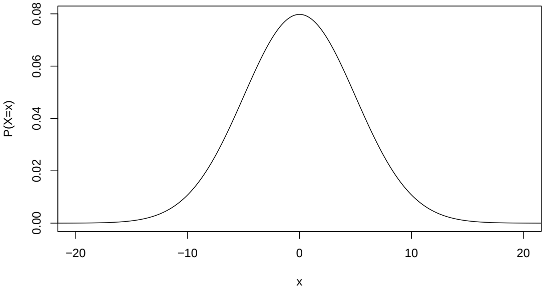
Intuitively Thinking: Measures of Spread

Consider the following (continuous) distribution with $\mu = 0$. Let's build a measure of expected "spread".



Intuitively Thinking: Measures of Spread

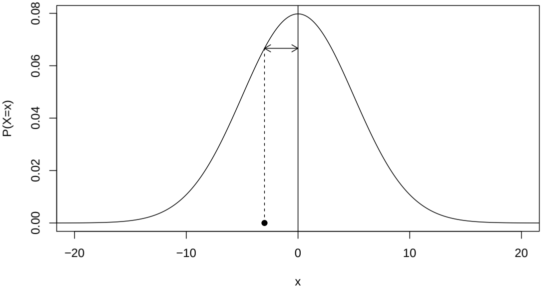
Consider the following (continuous) distribution with $\mu = 0$. Let's build a measure of expected "spread".



Let's define "spread" as the absolute deviation from μ: $|x - \mu|$.
i.e. $+$'ve & $-$'ve deviations are treated the same.

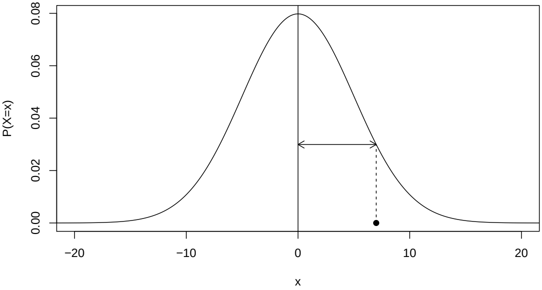
Intuitively Thinking: Measures of Spread

When $x = -3.0$, the abs. deviation from μ is $|-3.0 - \mu| = 3.0$.
Note $P(X = x) = 0.066$.



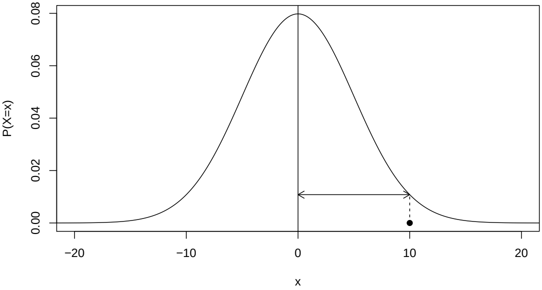
Intuitively Thinking: Measures of Spread

When $x = 7.0$, the abs. deviation from μ is $|7.0 - \mu| = 7.0$.
Note $P(X = x) = 0.030$.



Intuitively Thinking: Measures of Spread

When $x = 10.0$, the abs. deviation from μ is $|10.0 - \mu| = 10.0$.
Note $P(X = x) = 0.011$.



Intuitively Thinking: Measures of Spread

So say we do this for all x and take a weighted average of the $|x - \mu|$ where the weights are $P(X = x)$.

Voilà: Our notion of expected spread.

Variance

Estimators

Sample Mean as an Estimator

Bias

Recall from Earlier

One example of a non-representative sample is a biased sample. For example, convenience samples are samples where individuals who are easily accessible are more likely to be included.

Recall from Earlier

1. The Royal Air Force wants to study how resistant their airplanes are to bullets. They study the bullet holes on all the airplanes on the tarmac after an air battle against the Luftwaffe (German Air Force).
2. I want to know the average income of Reed graduates in the last 10 years. So I get the records of 10 randomly chosen Reedies. They all answer and I take the average.
3. Imagine it's 1993 i.e. almost all households have landlines. You want to know the average number of people in each household in Portland. You randomly pick out 500 phone numbers from the phone book and conduct a phone survey.
4. You want to know the prevalence of illegal downloading of TV shows among Reed students. You get the emails of 100 randomly chosen Reedies and ask them "How many times did you download a pirated TV show last week?"