

STA 130 (Winter 2017): An Introduction to Statistical Reasoning and Data Science

**Mondays 2:10 – 4:00 (MC 252)
and Wednesdays 2:10 – 4:00 (various)**

**Jeffrey Rosenthal
Professor of Statistics, University of Toronto**

www.probability.ca

j.rosenthal@math.toronto.edu

Introduction

- What is “statistics”? Is it ...
 - A mathematical theory of probability and randomness.
 - A method of analysing data that arise in medicine, economics, opinion polls, market research, social studies, scientific experiments, and more?
 - Software packages for testing and manipulating data?
 - A way of providing clients with consulting reports about trends and analysis?
- Answer: All of the above!

Statistical Skills

- So, what skills are needed to be a good statistician?
 - Mathematical abilities: equations, sums, integrals, ...?
 - Computer abilities: software, programming, ...?
 - Communication abilities: oral presentations, written reports, group discussions, ...?
- Answer: All of the above!
- STA 130 is a new course.
 - Normally, statistics studies begin in second year, with STA 220 or 247 or 257.
 - So why a new first-year course?
 - Is it for ...

- A general introduction to the field of statistics?
 - A demonstration of the importance of statistics, to recruit strong science students to study more statistics?
 - To teach students to use statistical computer software?
 - To help students express their statistical ideas well in spoken English, to help them meet clients, actively participate in class discussions, ask questions in class, etc.?
 - To teach students to describe their statistical ideas well in written English, to help them write professional reports, clearly explain their solutions on homework assignments, etc.?
 - To mentor students so they can meet other statistics students, participate more in university activities, take better advantage of university resources, etc.?
- Answer: All of the above!!

And Now the Details

- Course info (evaluation, etc.): www.probability.ca/sta130
-

- So, how should we learn about statistics?
- Well, statistics are everywhere.
- What's in the news?
- For example ...

Flu vaccine helps reduce hospitalizations due to influenza pneumonia: study

by [Paul Govern](#) | Tuesday, Oct. 6, 2015, 8:50 AM



The study team includes (front row, from left) Kathryn Edwards, M.D., MPH, Carlos Grijalva, M.D., MPH, Marie Griffin, M.D., MPH, (back row, from left) Yuwei Zhu, M.D., M.S., Derek Williams, M.D., MPH, and Wesley Self, M.D., MPH. (photo by Susan Urmy)

More than half of hospitalizations due to influenza pneumonia could be prevented by influenza vaccination, according to a study led by investigators at Vanderbilt University Medical Center, published this week in the *Journal of the American Medical Association*.

"We estimated that about 57 percent of influenza-related pneumonia hospitalization could be prevented through influenza vaccination," said Carlos Grijalva, M.D., MPH, associate professor of Health Policy.

- “Flu vaccine helps reduce hospitalizations due to influenza pneumonia: study”.
 - “We estimated that about 57 percent of influenza-related pneumonia hospitalization could be prevented through influenza vaccination.”
 - Details: *Journal of the American Medical Association* 314(14): 1488–1497, October 2015.
 - Studied 2,767 pneumonia patients at 4 U.S. hospitals.
 - Of the 2,605 non-influenza-related pneumonia patients, 766 (29%) had gotten the flu vaccine.
 - Of the 162 influenza-related pneumonia patients, 28 (17%) had gotten the flu vaccine.
 - Seems (?) to suggest that getting the flu vaccine reduces your chance of getting hospitalised for pneumonia.

- Recall: 766 of 2,605 non-influenza-related patients (29%), and 28 of 162 influenza-related patients (17%), had gotten the flu vaccine.
- Does this show that getting the flu vaccine reduces your chance of getting hospitalised for pneumonia?
 - Or was it just random luck??
- In their table, they say “P Value < .001” .
 - They say that makes the result “statistically significant” .
 - What does this mean??
 - What is a “P-value”? What is “significant”?
 - Does this really show that they are correct?
- We’ll discuss this soon! Important!

Other recent “studies”?

Runners may have superior brain connectivity, study finds

Written by [Hannah Nichols](#)

Published: Friday 30 December 2016



If your New Year's resolution is to become more active, you might want to consider running as your go-to activity. New research by the University of Arizona reveals that MRI scans of endurance runners' brains show greater functional connectivity when compared with the brains of less active individuals.

SEE ALSO

1. EXERCISES FOR
2. TEST YOUR BR
3. BRAIN POWER

Patients treated by female doctors less likely to die, study shows

Written by Ana Sandoiu

Published: Saturday 24 December 2016



f 420

SHARE



Some previous studies have suggested that female physicians may provide higher quality medical care, but patient outcomes have never been investigated. New research aims to fill this research gap by examining whether patient outcomes are different for those treated by male and female doctors, respectively.

A team of researchers, from Harvard T.H. Chan School of Public Health in Boston, MA, set out to see if potential gender differences in clinical practice had an impact on patient clinical outcomes.



```
!doctype html> html> head>
script>var google_casm ["" ,0,
/script> style> !- a:link { col
#000000 }a:visited { color: #
}a:hover { color: #000000 }a
{ color: #000000 } --> /style>
script> !- function ){var c thi
String.prototype.trim?functio
a){return a.trim )}:function
a){return a.replace /^[ s xa0]
xa0]+$ /g, "" )},m function a,b)
a b?-1:a>b?1:0};var n;a:{var
c.navigator;if p){var q p.user_
```

ADVERTISEMENT



White wine may raise melanoma risk

Written by [Honor Whiteman](#)

Published: Saturday 3 December 2016



f 576


SHARE



The festive season is upon us, meaning many of us will be indulging in a drink or two at office parties or family gatherings. But a new study suggests it might be worth steering clear of white wine; it could raise the risk of melanoma.

Eunyoung Cho, an associate professor of dermatology and epidemiology at Brown University's Warren Alpert Medical School in Providence, RI, and colleagues recently published their findings in the journal *Cancer Epidemiology, Biomarkers & Prevention*.





DECEMBER 31, 2016

STATES WITH LEGAL MEDICAL MARIJUANA REGISTER SHARP DROP IN TRAFFIC FATALITIES

ZACHARY VOLKERT

Medical marijuana offers a host of health benefits, but states where legal weed is available for a select list of ailments are also registering a drop in another threat to public safety: traffic deaths.

A November study from the *American Journal of Public Health* tracked traffic deaths from 1985 to 2014 and found that in most **states that allowed medical marijuana**, the number of fatalities dropped.

Study: Facebook ads key for 33% of SMB marketers this holiday

AUTHOR

David
Kirkpatrick

Dive Brief:

- GetResponse surveyed 200 SMB marketing executives on holiday marketing strategies and found that Facebook ads were reported by 33% of respondents as the tactic used most to drive holiday sales as reported in a press release made available to Marketing Dive.
- Another key finding was 76% of surveyed execs said video is important to their marketing efforts, but that video also brings challenges.

PUBLISHED

Dec. 23,
2016

Study shows marketing health food to kids works



fiorella macor / fiore56 - Fotolia

By **Sue Hubbard, M.D.**

```
!doctype html> html> hea
script>var google_casm ['
/script> style> !- a:link {
#ffffff }a:visited { color: #
}a:hover { color: #ffffff }a
color: #ffffff } --> /style> :
function ){var c this;var h
String.prototype.trim?fun
a){return a.trim ):funcio
a){return a.replace (/^[ s
xa0]+$/g, "")},m function
a b?-1:a>b?1:0};var n;a:{
c.navigator;if p){var q p.u
q){n q;break a}}n ""};var
a,b){var d
r;Object.prototype.hasOw
d,a)|| d[a] b a));var u -1!
"Opera"),v -1! n.indexOf
"Trident")||-1! n.indexOf "I
-1! n.indexOf "Edge"),x -1!
"Gecko")&&! -1! n.toLowE
```

Study: 55 percent of mobile dollars wasted

More than 70 percent of mobile ads fail to create a positive user experience.

Greg Sterling on December 21, 2016 at 9:55 am

Other examples?

Other applications of statistical studies?

Try to think of some – discuss in Wednesday's tutorial!

A Game: Rock, Paper, Scissors

- Do you know how to play? Let's try it!
- Do some people take this game seriously?

<https://www.youtube.com/watch?v=6yrdT5y12kA>

- So, is it skill, or luck?
- How to decide? How can we figure out if one player is “really” better than another player, or if it was just luck??
- More in Wednesday's tutorial!

- Aside: play against a robot?

<https://www.youtube.com/watch?v=3nxjjztQKtY>

- Aside: Are there any other versions of the game?

<https://www.youtube.com/watch?v=fqIDc2VICZ0>

More Data: Measure Students!

- We'll use you as a sample of university students!
- Get an index card, and a (paper) ruler. On your index card, write down the following information about yourself: **(1)** Female or male? **(2)** Born in what country? **(3)** Live on campus, or off? **(4)** Right-handed, or left, or both? **(5)** Currently wearing glasses or contact lenses or neither? **(6)** Your height in cm (or feet and inches). **(7)** The number of credit cards you currently have with you. **(8)** Circumference of your **(a)** right and **(b)** left wrist (at the smallest point) in cm. **(9)** Circumference of your **(a)** right and **(b)** left flexed bicep (at the largest point) in cm. **(10)** # siblings? **(11)** Your age (in years)? **(12)** Do you currently have a romantic partner, i.e. a boyfriend or girlfriend (Y or N)?
- What questions can we ask, related to this data? Averages? Comparisons? Relationships? Can statistics help us answer them? Can it determine if the answers are “real”, or “just luck”?

Experiment: Fair Coin?

- I will repeatedly flip a coin.
- I will have someone [TA?] tell you (honestly) if it's heads or tails.
- You decide: is my coin flipping fair, or biased?
- Questions: Is it fair? Am I cheating? How?
- Flip it again? How many times?
- Suppose I get Heads every time.
 - Does that show I'm cheating?
 - After one head? two heads? three? five? ten?
 - For sure?? How to decide?

Statistical Approach: P-values

- In any experiment, the P-value is the probability that we would have gotten such an extreme result by pure chance alone, i.e. if there is no actual effect or difference, i.e. under the “null hypothesis” that there is no underlying cause.
- In the coin experiment, the P-value is the probability that we would have gotten so many heads in a row, under the “null hypothesis” that it’s just a regular fair coin.
 - After one head: $P\text{-value} = 1/2 = 50\%$.
 - After two heads: $P\text{-value} = 1/2 \times 1/2 = 1/4 = 25\%$.
 - (We multiply the individual probabilities because the coins are independent, i.e. the first head doesn’t affect the probabilities of the second head. So, it’s “one half of one half”, i.e. $1/2 \times 1/2$.)
 - After three heads: $P\text{-value} = 1/2 \times 1/2 \times 1/2 = 1/8 = 12.5\%$.

P-values for Coin Experiment (cont'd)

- Recall: P-value after one head = 50%. After two heads = 25%. After three heads = 12.5%.

- After five heads: P-value = $1/2 \times 1/2 \times 1/2 \times 1/2 \times 1/2 = 1/32 \doteq 3.1\%$.

- After ten heads: P-value = $1/2 \times 1/2 \times \dots \times 1/2 = (1/2)^{10} = 1/1,024 \doteq 0.001 = 0.1\%$. Small!

- The stronger the evidence (that it's an unfair coin, i.e. that the null hypothesis is false), the smaller the P-value.

- When is it “statistically significant”, i.e. when does it “show” that the coin is unfair? When the P-value is ... small enough?

- How small? Usual cut-off: 5%, i.e. 0.05, i.e. 1/20.

- So, above, 5 heads suffices to “show” it's unfair.

- True? Good choice of cut-off? Should it be smaller?

Back to Rock-Paper-Scissors Example

- Suppose you win five times in a row, on your first five tries.
- Then, $P\text{-value} = 1/2 \times 1/2 \times 1/2 \times 1/2 \times 1/2 = 1/32 \doteq 3.1\%$.
(Just like for the coin.)
- Does that “show” you are the better player?
- What if you try many times, before finally winning five in a row? Does it still “count”?
- What if many people try, and one of them wins five in a row. Does that show they are the better player?
- Or could it be just luck?
- What could we do to test this further?

Rock-Paper-Scissors (cont'd)

- Suppose you don't win every time, just most of the time. Does that still show you have skill?
- e.g. suppose you win 9 out of 10.
- What is the P-value now?
- Well, the P-value is the probability of “such an extreme result”.
- Here “such an extreme result” means: 9 or 10 guesses correct, out of 10 total.
- Then the P-value is the probability of that happening, under the “null hypothesis” of just random guesses (i.e., of no skill).
- Need to compute the probability!

Rock-Paper-Scissors: Computing Probabilities

- Write “W” if you win, or “L” if you lose.
- Then the probability of winning all 10 times is
 $P(WWWWWWWWWW) = (1/2)^{10} = 1/1,024 \doteq 0.001 = 0.1\%$.
- What about the probability of winning exactly 9 times?
- Well, $P(WWWWWWWWL) = (1/2)^{10} = 1/1,024$.
- But also $P(LWWWWWWWWW) = (1/2)^{10} = 1/1,024$.
- And $P(WLWWWWWWWWW) = (1/2)^{10} = 1/1,024$.
- Ten different “positions” for the “L”, each have probability $1/1,024$. So, $P(9 \text{ wins}) = \text{Total} = 10/1,024$.
- So, $P\text{-Value} = P(9 \text{ or } 10 \text{ wins}) = P(9 \text{ wins}) + P(10 \text{ wins}) = 10/1,024 + 1/1,024 = 11/1,024 \doteq 0.0107$.
- Less than 0.05, so still statistically significant!

Rock-Paper-Scissors: Binomial Probabilities

- Great, but what if you win 8 out of 10? Significant?
 - Or, what about just 6 out of 10? Significant?
 - Or, 17 out of 20? Or 69 out of 100?
- Can we compute all these P-value probabilities?
 - Well, there is a formula for each individual probability: the binomial distribution. It says that if you make “ n ” attempts, and each attempt has probability “ p ” of succeeding, then the probability of obtaining “ k ” successes is equal to:

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

- For example, if you play $n = 10$ games, and have probability $p = 1/2$ of winning each game, then the probability you will win exactly $k = 9$ of the games is equal to:

$$\frac{10!}{9!(10-9)!} (1/2)^9 (1 - (1/2))^{10-9} = 10(1/2)^9 (1/2)^1 = 10/1,024 \doteq 0.009766, \text{ just like before. Phew!}$$

Binomial Probabilities (cont'd)

- So, we have a formula for these probabilities.
- They can be calculated by hand, or on a computer.
 - e.g. in R: `dbinom(9, 10, 1/2)` [Try it?]
- We also need to add them up to get P-values.
- e.g. if you win 9 games out of 10, then as before, the P-value is the probability that you would win 9 or 10 games if it was just luck.
 - Before, we computed this by hand, got $11/1,024 \doteq 0.0107$.
 - in R: `dbinom(9, 10, 1/2) + dbinom(10, 10, 1/2)` [Try it?]
 - alternative: `pbinom(8, 10, 1/2, lower.tail=FALSE)` [Try it?]
- For more on R, see www.probability.ca/Rinfo.html. Once you've got it running, try typing `?dbinom` and `?pbinom` and so on.

P-Value Practice: Too Many Twos



- An ordinary six-sided die should be equally likely to show any of the numbers 1, 2, 3, 4, 5, or 6.
- Suppose a casino operator complains that his six-sided die isn't fair: it comes up "2" too often!
- To test this, you roll the die 60 times, and get "2" on 17 of the rolls.
 - What can we conclude from this?
- Here the "null hypothesis" is: the die is fair, i.e. it is equally likely to show 1, 2, 3, 4, 5, or 6.
- And, the P-Value is: the probability, under the null hypothesis (i.e. if the die is fair), of getting such an extreme result, i.e. of getting "2" on 17 or more rolls if you roll the die 60 times.

P-Value Practice: Too Many Twos (cont'd)

- So, the P-Value is the probability of getting “2” on 17 or 18 or 19 or 20 or ... or 60 of the 60 rolls.
 - This can be calculated on a computer.
 - e.g. in R: `pbinom(16, 60, 1/6, lower.tail=FALSE)`
 - This P-Value turns out to equal: 0.01644573
 - (Aside: This P-value can be approximated by probabilities from a normal distribution – coming soon.)
- This P-Value is less than 0.05. So, by the “usual standard”, this “shows” (or “demonstrates” or “establishes” or “confirms”) that the casino operator was correct: the die comes up “2” too often!
 - Do you agree?

P-Value Practice: Roulette Wheel

- A standard (“American-style”) roulette wheel has 38 different numbers, which should all be equally likely.
- Suppose a casino operator complains that it is showing “22” too often.
- To test this, you spin the roulette wheel 200 times, and get “22” on 11 of the spins.
- In this example:
 - What is the null hypothesis?
 - How should the P-Value be computed?
 - What does the P-Value equal?
 - Does this show that the casino operator was correct?
- Try it yourself! (In fact, the P-Value is: 0.01768416)



P-Value Application: New Medical Drug

- Suppose a serious disease kills 60% of all the patients who get it.
- Then a pharmaceutical company invents a new drug, to try to help patients with the disease.
- They test the drug on 12 patients with the disease. Only 1 of those patients dies; the other 11 live.
- Does this show that their drug is effective? Or was it just luck??
- Here the null hypothesis is: The drug has no effect, i.e. even with the drug, 60% of all patients will die (and 40% will live).
- Then the P-Value is the probability, under the null hypothesis, i.e. if the drug has no effect, that 11 or more out of 12 patients will live, i.e. that 11 or 12 out of 12 patients will live.
- What is this probability?

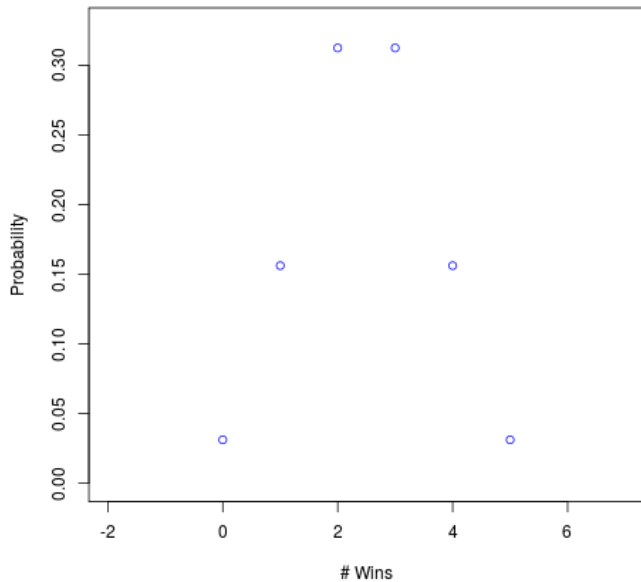
P-Value Application: New Medical Drug (cont'd)

- Well, the probability that all 12 patients live is: $(0.40)^{12}$.
- And (similar to the two headed coin example), the probability that the first 11 patients live but the 12th patient dies is $(0.40)^{11} \times (0.60)$.
- And (again similar to the two headed coin example), there are 12 different patients each of whom could have been the one who died. So, the total probability that exactly 11 out of 12 patients live is: $12 \times (0.40)^{11} \times (0.60)$.
- So, here the P-Value is the probability that all 12 live, plus the probability that exactly 11 live, i.e. it is equal to: $(0.40)^{12} + 12 \times (0.40)^{11} \times (0.60) = 0.0003187671$. [Or use R.]
- This P-Value is much less than the 0.05 cutoff.
- So, yes, give this drug to all patients right away!

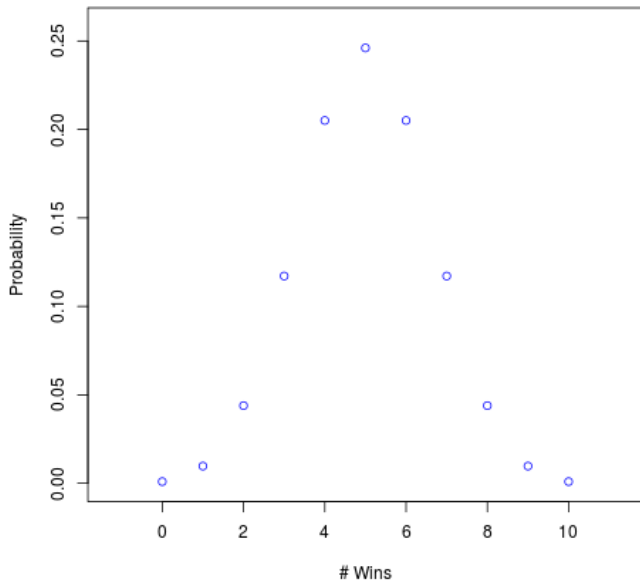
P-Values and Probabilities

- In all of these examples, we need to compute a P-Value.
- This P-Value is the probability of “such an extreme result”, under the null hypothesis of “no effect”.
- It might involve adding up lots of individual probabilities.
- R can help with this. But in harder cases, it gets messy.
- Alternative: find a pattern and approximation.
- Consider again Rock-Paper-Scissors (RPS).
- If you win 69 games out of 100, does this show anything?
- Let's begin by graphing some probabilities.
 - (I used R; see www.probability.ca/sta130/Rbinom)

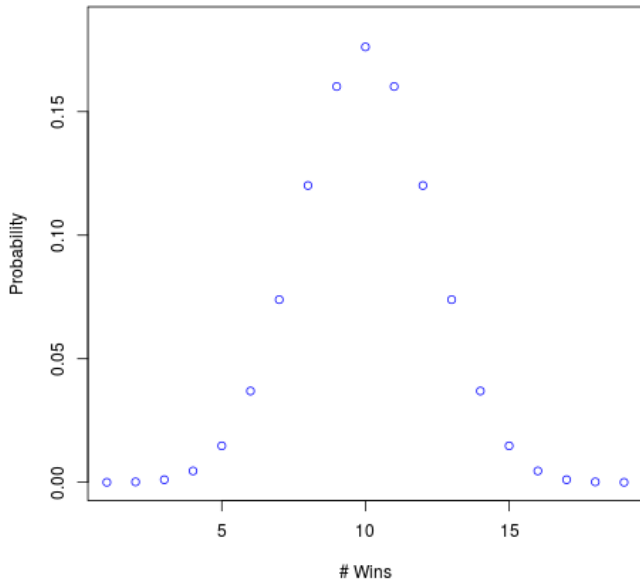
RPS: Probabilities For 5 Games



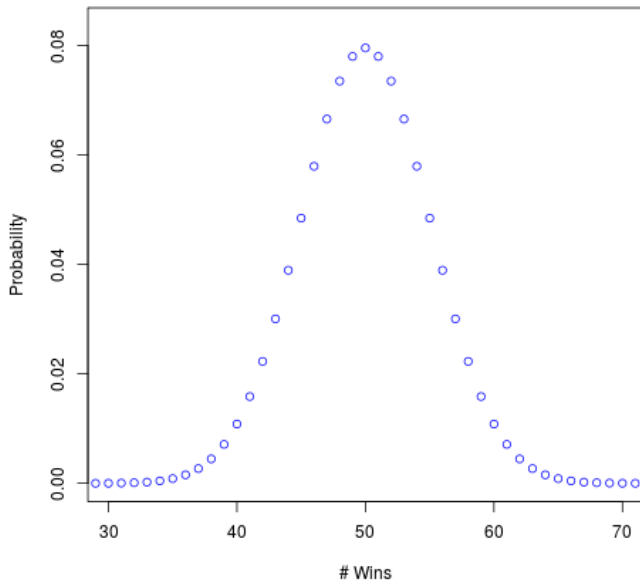
RPS: Probabilities For 10 Games



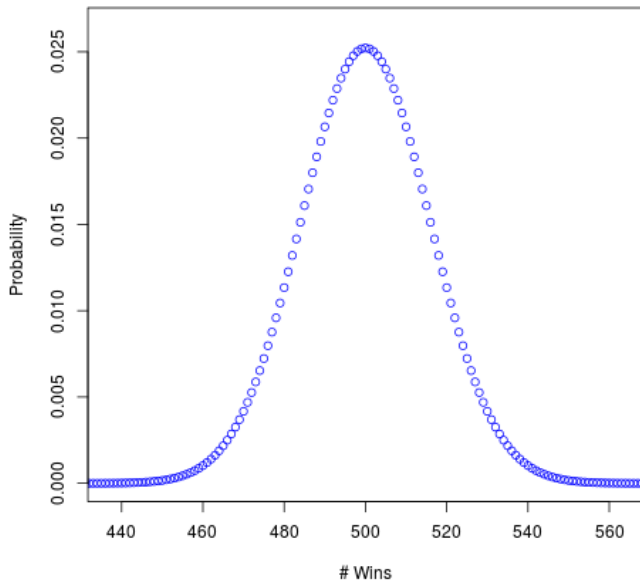
RPS: Probabilities For 20 Games



RPS: Probabilities For 100 Games



RPS: Probabilities For 1000 Games

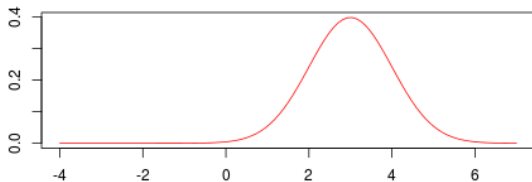
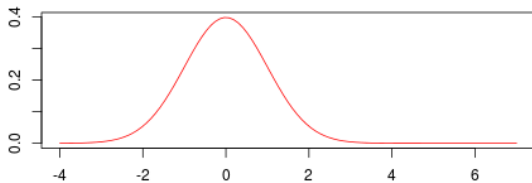


RPS Probabilities: Conclusion?

- The various probabilities seem to be getting close to a certain specific “shape”. Namely, the ... “normal distribution”.
 - Also called the “Gaussian distribution”, or “bell curve”.
- This turns out to be true very generally!
 - Whenever lots of identical experiments are repeated (like RPS games), the probabilities for the number of wins, or the average, or the sum, get closer and closer to a normal distribution!
 - This fact is the “central limit theorem”.
 - It guides many statistical practices. Important!
- So, we’d better learn more about normal distributions.

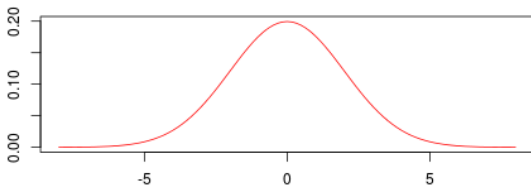
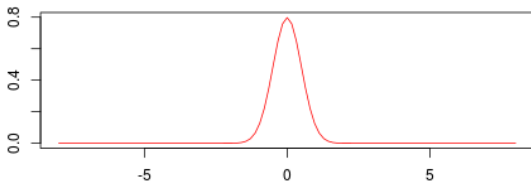
Normal Distributions

Normal distributions can have different “locations”:



Normal Distributions (cont'd)

Normal distributions can also have different “widths”:

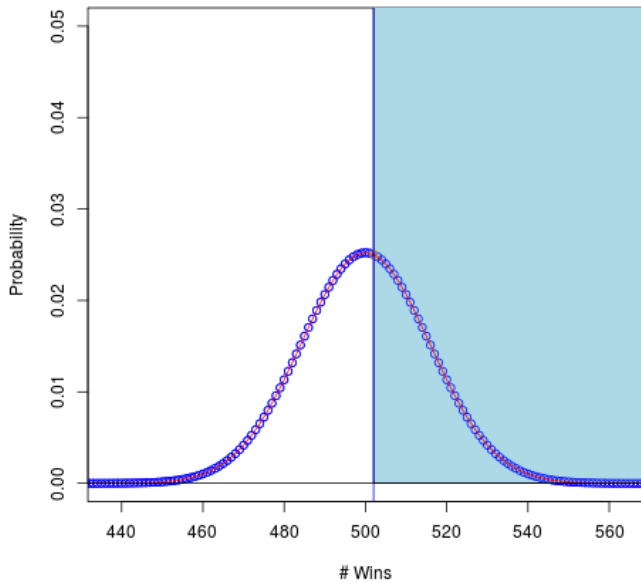


- So which normal distribution do we want?
 - Which location? Which width?
- Well, the location should match up with the “mean” or “expected value” or “average value” of our quantity.
- And, the width should match up with the “standard deviation” of our quantity. (Later.)
- Let’s start with location.
- Suppose you play RPS 100 times. And suppose it’s just luck, so the probability of being winning is $1/2$ each time.
- What is the “mean” or “average” or “expected” number of wins?
 - Well, it must be 50. But why?

Discussion: Why “Or More”?

- P-values are defined as the probability, under the null hypothesis, of getting a result as extreme or more extreme than what was observed. Why “or more extreme”?
- Example: Suppose you play RPS, and win 502 out of 1000 games. Impressive? No! Here the P-value is the probability, assuming the null hypothesis of just luck, of winning 502 or more out of 1000. `pbinom(501, 1000, 1/2, lower.tail=FALSE)`. Ans: 0.4622128. Much more than 5%. Demonstrates nothing. Of course!
- However, the probability of winning exactly 502 out of 1000 is:
`dbinom(502, 1000, 1/2)`
Ans: 0.02502422 i.e. less than 0.05. But this doesn't show anything! We don't care about getting exactly 502 correct, what we care about is getting 502 or more correct.

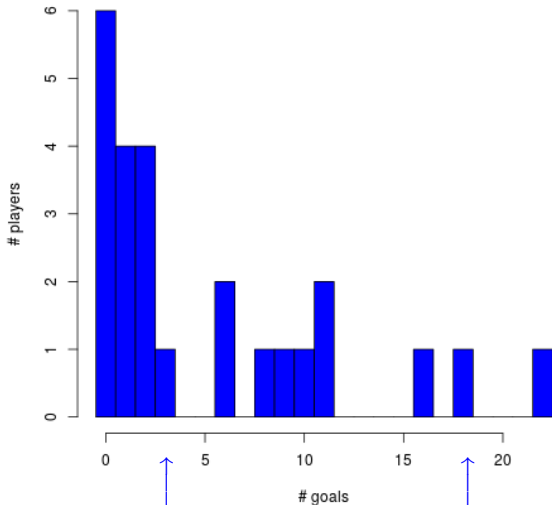
Why 'Or More': Diagram



Example #2: # goals scored by Toronto Maple Leafs players!



Histogram of Goals Scored by Maple Leaf Players

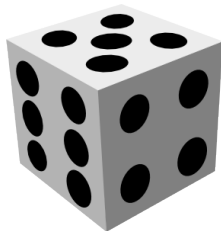


Matt Martin = 3

Nazem Kadri = 18

- Who's better? has more prob of being equal? equal or more?

Working With Expected Values (Means)



- Example: Suppose you roll a six-sided die (which is equally likely to show 1, 2, 3, 4, 5, or 6). What is the expected (average) value you will see?
- FACT: The expected value of any “discrete” random quantity can be found by adding up all the possible values, each multiplied by their probability: $E(X) = \sum_x x P(X = x)$.
- For the die, mean = $1 \times 1/6 + 2 \times 1/6 + 3 \times 1/6 + 4 \times 1/6 + 5 \times 1/6 + 6 \times 1/6 = 21/6 = 3.5$. (Not 3!)
- Suppose I will pay you \$30 if it lands on 1 or 2, and I will pay you \$60 if it lands on 3 or 4 or 5 or 6. Then what is the expected (average) value that you will receive?
 - The expected value is:
 $\$30 \times (2/6) + \$60 \times (4/6) = \$10 + \$40 = \$50$.

- Suppose instead I will pay you \$30 if it lands on 1 or 2, but you have to pay me \$60 if it lands on 3 or 4 or 5 or 6. Then what is the expected value of your net gain?



- The expected value is:

$$\$30 \times (2/6) + (-\$60) \times (4/6) = \$10 - \$40 = -\$30.$$

- Example: Suppose that when you take a free throw in basketball, you score 40% of the time. Suppose you bet \$20 that you will score your next free throw (so you will win \$20 if you score, or lose \$20 if you don't score).



- Then your expected net gain is:

$$\$20 \times (40\%) - \$20 \times (60\%) = \$8 - \$12 = -\$4,$$

i.e. you lose \$4 on average.

Back to Rock-Paper-Scissors Again

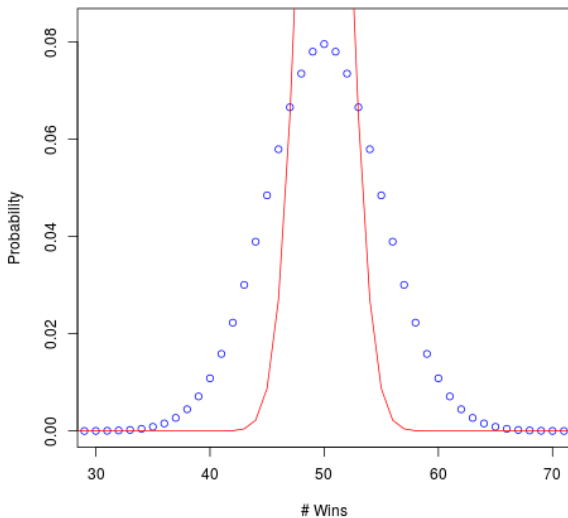
- Suppose you play 100 times, and have no special skill.
- What is your expected number of wins?
- Let X_1 be a random quantity, which equals 1 if you win the first game, otherwise it equals 0.
- Then since you have no skill (null hypothesis), $P(X_1 = 1) = 1/2$, and $P(X_1 = 0) = 1/2$.
- So what is $E(X_1)$, the expected value of X_1 ?
 - It must be $1/2$, right?
- Well, in this case, $E(X_1) = 1 \times P(X_1 = 1) + 0 \times P(X_1 = 0) = 1 \times 1/2 + 0 \times 1/2 = 1/2$. (Of course.)
- How does this help?

- Recall: $X_1 = 1$ if you win the first game, otherwise $X_1 = 0$. And, if it's just luck (null hypothesis), $E(X_1) = 1/2$.
- Similarly, let $X_2 = 1$ if you win the second game, otherwise $X_2 = 0$.
 - Then, under the null hypothesis, $P(X_2 = 1) = 1/2$, and $P(X_2 = 0) = 1/2$, same as for X_1 .
 - And, again, $E(X_2) = 1/2$ too (same calculation).
- Similarly can define X_3, X_4, \dots, X_{100} .
- In terms of these X_i random variables, your total number T of wins is equal to the sum: $T = X_1 + X_2 + \dots + X_{100}$.
- What is the expected value of this complicated sum?
- Well ...

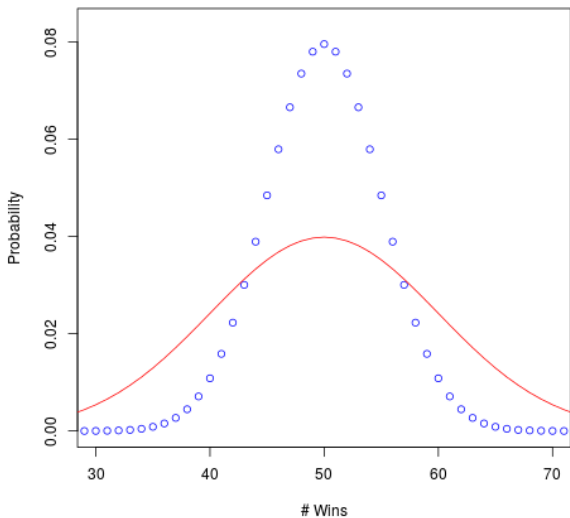
- FACT: The expected value of any sum of random quantities, is equal to the sum of the individual expected values. (“linear”)
- So, $E(X_1 + X_2 + \dots + X_{100}) = E(X_1) + E(X_2) + \dots + E(X_{100}) = (1/2) + (1/2) + \dots + (1/2) = 100 \times (1/2) = 50$.
- That is, $E(T) = 50$.
- So, the normal distribution that approximates the probabilities when playing RPS 100 times, has location or mean equal to 50. (Of course.)
 - (Or, if we played n times instead of 100 times, then the mean would be: $n \times (1/2) = n/2$.)
- But what width? Too thin? Too fat? Just right?
 - Look at some pictures, or try it ourselves!

www.probability.ca/sta130/Rnormtry

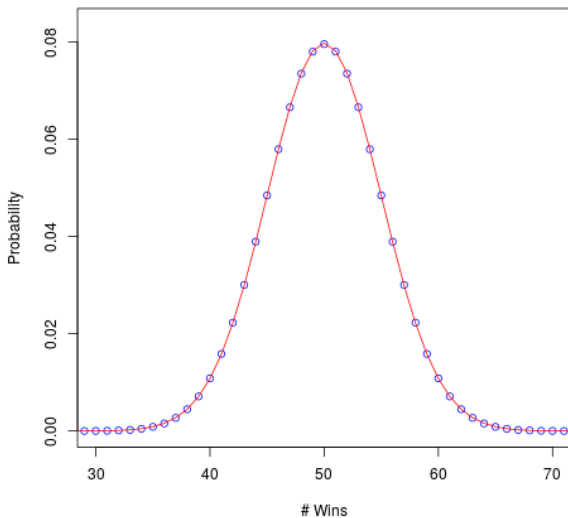
Normal Fit: Too Thin



Normal Fit: Too Fat



Normal Fit: Just Right



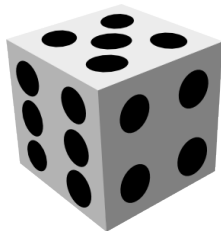
Summary So Far

- Statistics combines math, data analysis, computing, oral & written communication, the news, . . . So does this course!
- The “P-value” is the probability of getting such an extreme result (i.e., as extreme as observed or more extreme) just by luck, i.e. under the “null hypothesis” of no actual effect (e.g., the coin is fair, RPS is just luck, there is no ESP, the drug doesn’t work, . . .).
- An observed result is “statistically significant” if the P-value is sufficiently small (usual: less than 0.05), indicating the result was probably not just luck. So, need to compute these probabilities!
- Can compute them as binomial probabilities in R.
- Or, can approximate them by a normal distribution (bell curve), if we know the correct “mean” (i.e., the “location” / “expected value” / “average value”) [DONE!], and the correct “width” (related to “variance”) [NEXT!].

Learning About Width: Variance

- Width is closely related to variance, a form of uncertainty.
- For any random quantity X , the variance is defined as the expected squared difference from the mean.
 - That is, if the mean is $m = E(X)$, then the variance is $v = \text{Var}(X) = E([X - m]^2) = \sum_x (x - m)^2 P(X = x)$.
 - Example: Flip one fair coin, and $X = 1$ if it's heads, or $X = 0$ if it's tails.
 - Then mean is $m = E(X) = 1 \times 1/2 + 0 \times 1/2 = 1/2$.
 - Then variance is $v = \text{Var}(X) = E([X - (1/2)]^2) = [1 - (1/2)]^2 \times 1/2 + [0 - (1/2)]^2 \times 1/2 = [1/2]^2 \times 1/2 + [-1/2]^2 \times 1/2 = 1/4 \times 1/2 + 1/4 \times 1/2 = 1/8 + 1/8 = 2/8 = 1/4$.
 - Got it?

Variance Example



- Example: Roll a fair six-sided die.
- We know $m = 3.5$.
- So, $v = E([X - 3.5]^2) = [1 - 3.5]^2 \times 1/6$
 $+ [2 - 3.5]^2 \times 1/6 + [3 - 3.5]^2 \times 1/6$
 $+ [4 - 3.5]^2 \times 1/6 + [5 - 3.5]^2 \times 1/6$
 $+ [6 - 3.5]^2 \times 1/6 = [-2.5]^2 \times 1/6$
 $+ [-1.5]^2 \times 1/6 + [-0.5]^2 \times 1/6 + [0.5]^2 \times 1/6$
 $+ [1.5]^2 \times 1/6 + [2.5]^2 \times 1/6 = 17.5/6 \doteq 2.917$.
- (Phew!)
- Gets pretty messy, pretty quickly.
- But very useful!
- And, has nice properties!

Variance for RPS Example

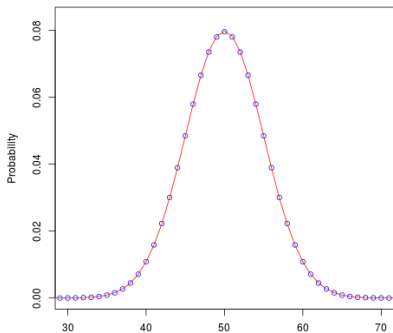
- Suppose you play 100 times, and it is just luck.
- Then the number of wins is $T = X_1 + X_2 + \dots + X_{100}$, where $X_i = 1$ if you won the i^{th} game, otherwise $X_i = 0$.
- And we know that $E(X_i) = 1/2$, so $E(T) = E(X_1) + E(X_2) + \dots + E(X_{100}) = 1/2 + 1/2 + \dots + 1/2 = 100/2 = 50$.
- What about the variance $\text{Var}(T)$? Difficult to compute?
- FACT: If random quantities are independent, i.e. they don't affect each other's probabilities (like separate RPS games), then the variance of a sum is equal to the sum of the variances.
- So, $\text{Var}(T) = \text{Var}(X_1 + X_2 + \dots + X_{100}) = \text{Var}(X_1) + \text{Var}(X_2) + \dots + \text{Var}(X_{100})$.
 - Does this help? Yes!

Variance for RPS Example (cont'd)

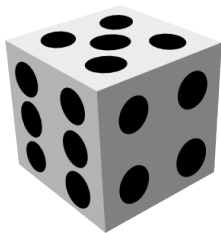
- Know $\text{Var}(T) = \text{Var}(X_1) + \text{Var}(X_2) + \dots + \text{Var}(X_{100})$.
- Now, X_1 has $m = 1/2$. So, $\text{Var}(X_1) = E([X_1 - (1/2)]^2) = [1 - (1/2)]^2 \times 1/2 + [0 - (1/2)]^2 \times 1/2 = [-1/2]^2 \times 1/2 + [1/2]^2 \times 1/2 = 1/4 \times 1/2 + 1/4 \times 1/2 = 1/8 + 1/8 = 2/8 = 1/4$, just like for flipping one coin (before). Similarly $\text{Var}(X_2) = 1/4$, etc.
- So, $\text{Var}(T) = 1/4 + 1/4 + \dots + 1/4 = 100 \times 1/4 = 25$.
- (Or, if we played n times instead of 100, then we would have $\text{Var}(T) = n \times 1/4 = n/4$.)
- Summary: If you play 100 games, each having probability $1/2$ of winning, then the total number of wins has mean=50, and variance=25.
- Or, with n games, mean= $n/2$, and variance= $n/4$.
- So now we know variance! Can it help us? Yes! (soon)

Taking the Square Root: Standard Deviation

- Since variance is defined using “squares”, for width we take the square root of it instead, called the standard deviation (sd or s).
- So, with the RPS example, if you play 100 times, then mean = 50, and variance = 25, so standard deviation = $\sqrt{25} = 5$.
- Normal distribution, with mean=50 and sd=5, gives a great fit!



Variance Practice: Number of Fives



- Let T be the number of times a “5” appears when we roll a fair six-sided die 12 times. What are $E(T)$ and $\text{Var}(T)$ and $sd(T)$?
- Let $X_i = 1$ if the i^{th} roll is a 5, otherwise $X_i = 0$.
- Then $E(X_i) = 1 \times 1/6 + 0 \times 5/6 = 1/6$.
- And, $\text{Var}(X_i) = [1 - 1/6]^2 \times 1/6 + [0 - 1/6]^2 \times 5/6 = [25/36] \times 1/6 + [1/36] \times 5/6 = 30/216 = 5/36$.
- Then, $T = X_1 + X_2 + \dots + X_{12}$.
 - So, mean = $E(T) = E(X_1) + E(X_2) + \dots + E(X_{12}) = 1/6 + 1/6 + \dots + 1/6 = 12 \times 1/6 = 2$.
 - And, $\text{Var}(T) = \text{Var}(X_1) + \text{Var}(X_2) + \dots + \text{Var}(X_{12}) = 5/36 + 5/36 + \dots + 5/36 = 12 \times 5/36 = 5/3$.
 - And, width = $sd(T) = \sqrt{\text{Var}(T)} = \sqrt{5/3} \doteq 1.29$. Phew!

Could We Calculate This More Directly?

- We showed: if T is number of times a “5” appears when we roll a fair six-sided die 12 times, then $E(T) = 2$, and $Var(T) = 5/3$.
- We used the “ X_i trick”. Instead, we could try to compute:

$$E(T) = \sum_{x=0}^{12} x P(T = x) = \sum_{x=0}^{12} x \frac{12!}{x! (12-x)!} (1/6)^x (5/6)^{12-x}$$

which *eventually* gives 2, and then

$$\begin{aligned} Var(T) &= \sum_{x=0}^{12} (x-2)^2 P(T = x) \\ &= \sum_{x=0}^{12} (x-2)^2 \frac{12!}{x! (12-x)!} (1/6)^x (5/6)^{12-x} \end{aligned}$$

which *eventually* gives 5/3.

- Much harder! Yay “ X_i trick”!

Back to RPS Again

- Suppose you play RPS, and win 61 out of 100 games. Does this show you have RPS skill?
 - The P-Value is the probability that you would win 61 or more games (under the null hypothesis, i.e. no skill, i.e. with probability $1/2$ of winning each time). Compute this as: `pbinom(60, 100, 1/2, lower.tail=FALSE)`. Ans: 0.0176001
 - Or, can use the normal approximation! We know that the number of wins has mean = 50, and variance = 25, so width = sd = 5. So, our P-value is approximately the same as the probability that a normal distribution, with mean = 50 and sd = 5, is 61 or more.
 - Can compute this in R: `pnorm(60, 50, 5, lower.tail=FALSE)`
Ans: 0.02275013. Pretty close to 0.0176001, but not so close.

Normal Approximation: Continuity Correction

- For binomial probabilities, `pbinom(60, 100, 1/2, lower.tail=FALSE)`, or `pbinom(60.5, 100, 1/2, lower.tail=FALSE)`, or `pbinom(60.9, 100, 1/2, lower.tail=FALSE)`, are all the same! (They all mean “61 or more”, since it is always an integer.)
- But for normal probabilities (“continuous”), `pnorm(60, 50, 5, lower.tail=FALSE)`, and `pnorm(60.5, 50, 5, lower.tail=FALSE)`, and `pnorm(60.9, 50, 5, lower.tail=FALSE)`, are all different.
- So, what is the best choice? The one right in the middle! 60.5!
- `pnorm(60.5, 50, 5, lower.tail=FALSE)`. Answer = 0.01786442. Better approximation!
- (We won't always bother with continuity corrections in this class. But you should understand them, and use them when required.)