<u>Monday, September 14</u>: Unit 3: Independence

Are the outcomes of consecutive free throws independent or is it possible to get "in the zone?" In other words, is the *Ability* to make a free throw higher following a made free throw than following a missed free throw?

Def: If the *Ability* to make a shot does NOT depend on the outcome of previous shots, then the shots are <u>independent</u>.

Here is some data from a player who took 30 pairs of free throws.

	Made the first	Missed the first	Total
Made the second	9	6	15
Missed the second	7	8	15
Total	16	14	30

Given that she made the first shot, what percent of the time did she make the second shot?

P(made the second | made the first) =

Given that she missed the first shot, what percent of time did she make the second shot?

P(made the second | missed the first) =

What is the difference in these percentages?

Does this difference in *Performance* give convincing evidence that her *Ability* to make a free throw is higher following a made free throw than following a missed free throw? Or could the difference be due to *Random Chance*?

State the hypotheses we are interested in testing:

H_0: The player's *Ability* to make a free throw is the same following a made shot and following a missed shot.

H_a: The player's *Ability* to make a free throw is higher following a made shot than following a missed shot.

How likely is it to get a difference in *Performance* of 13.4% just by *Random Chance*? In other words, what is the *p*-value?

Simulation using index cards: Same as Unit 2!

HW #9: Unit 3, SR (1-3), Problems 1-3

<u>Tuesday/Wednesday, September 15/16</u>: Unit 3: Streakiness

Here are the outcomes for the 2008 Green Bay Packers:

W W L L L W W L L W L L L L L W

Based on their *Performance*, can we conclude that the Packer's *Ability* to win is greater following a win than following a loss?

To investigate, let's look at each pair of games and make a two-way table:

	Won the first	Lost the first	Total
Won the second			
Lost the second			
Total			

P(won the second | won the first) =

P(won the second | lost the first) =

Difference =

Does this difference in *Performance* give convincing evidence that their *Ability* to win is higher following win than following a loss? Or could the difference be due to *Random Chance*? State the hypotheses, interpret the *p*-value, and make an appropriate conclusion.

Another way to investigate independence is to look at streaks of wins or losses. If a team is more likely to win after a win (and also more likely to lose after a loss), we would expect a team to have some long streaks of wins or losses during the season.

How long was Green Bay's longest streak?

Does this *Performance* streak give evidence that the Packers *Ability* to win is higher following a win or could it have been *Random Chance*? State the hypotheses, design and conduct a simulation, interpret the *p*-value and make an appropriate conclusion.

HW #10: Unit 3, SR (3-7) Problems (4-5)

<u>Thursday/Friday, September 17-18</u>: Unit 3: More streakiness

In the 2008-2009 season, the LA Clippers had an overall record of 19-63 and a losing streak of 12 consecutive games in December and January.

Does this streak of bad *Performances* give evidence that the Clipper's *Ability* to win is lower following a loss or could it have been *Random Chance*? State the hypotheses, design and conduct a simulation, interpret the *p*-value and make an appropriate conclusion.

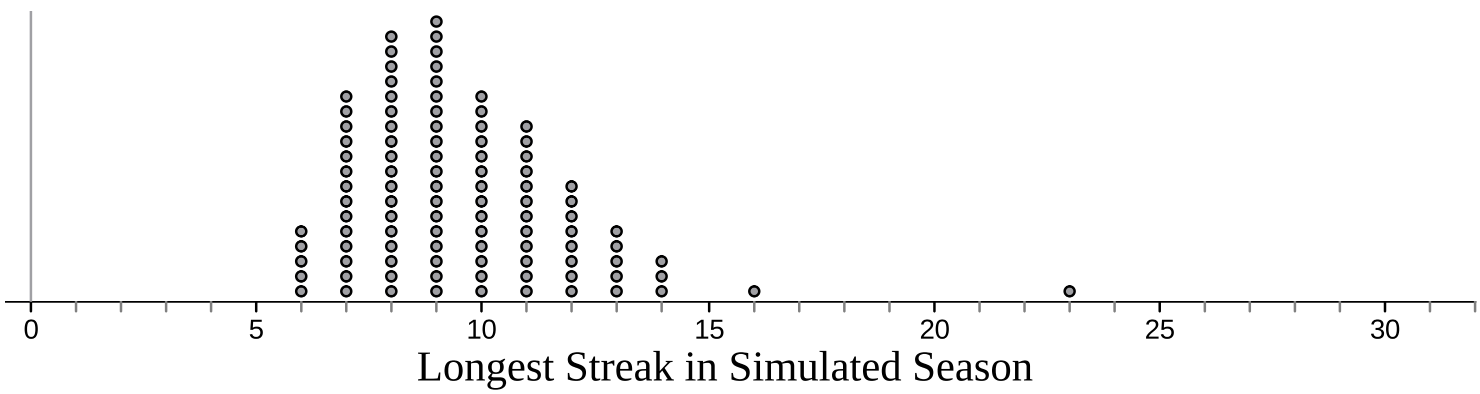
HW #11: Unit 3, Problems (6-7)

<u>Monday, September 21</u>: Unit 3: Type I and Type II errors

The 2002 Oakland A's were 103-59 with an incredible 20 game win streak. Does this streak of good *Performances* give evidence that the A's *Ability* to win is higher following a win than following a loss, or could it have been *Random Chance*?

State the hypotheses we are interested in testing.

A simulation was conducted assuming that the outcome of A's games are independent. In each trial, the longest streak was recorded. Based on the results below, interpret the *p*-value and make an appropriate conclusion.



Of course, it is always possible that the decision we make is wrong.

Type I error: Deciding that the alternative hypothesis is true when in reality the null hypothesis is true.

Sometimes the *p*-value will indicate that *Random Chance* isn't a likely explanation when in reality it was just *Random Chance*.

How often will this happen?

If the null hypothesis is true and we reject it whenever the *p*-value is less than .05, we will make a Type I error 5% of the time.

Type II error: Not deciding that the alternative hypothesis is true when in reality it is true.

Sometimes the *p*-value isn't small enough to rule out *Random Chance* when in reality it wasn't *Random Chance*.

How often will this happen?

It depends on how big the difference in *Ability* is. If the two *Ability* levels are very close, we will make lots of Type II errors. However, we are less likely to make a Type II error if we have lots of *Performances* on which to base our decision.

HW #12: Unit 3, SR (8-9) Problems (8-10)

<u>Tuesday/Wednesday, September 22/23</u>: Computer lab/work on projects

<u>Thursday/Friday, September 24/25</u>: Review/work on projects

Type I errors and multiple tests:

Assume that the outcomes of games are independent for all 32 NFL teams. If we did a test for one team, what is the probability we make a Type I error?

This means that for about 1 out of every 20 teams we will make a Type I error—that is, deciding that they have a greater *Ability* to win following a win when in reality it was just by *Random Chance*.

What is the probability we <u>avoid</u> a Type I error for a particular team?

What is the probability we <u>avoid</u> a Type I error for <u>all</u> 32 teams?

This means that if we did a test for all 32 teams and all the outcomes were independent, we have an 81% chance of mistakenly concluding that at least one team has a greater *Ability* to win following a win.

Lesson: If you look at enough players/teams, you will find some with *Performances* that are unlikely to happen by chance. People in the media often are guilty of "cherry-picking" these performances to illustrate something that may be entirely due to chance.

Connections: There is still a big debate about the issue of independence. However, in the previous two units and in future units we will be assuming that the outcomes of trials in sports are independent.

<u>Monday, September 28</u>: Test/Projects due