Statistics in Sports (Math 192) - Quiz 5
Spring 2009 - Brad Hartlaub

Name Brad - Solutions

Directions: Please answer all of the questions below. The point values for each problem are indicated in parentheses. Partial credit will be awarded if you show your work. Be careful not to spend too much time on any one part. You may not use any notes or the text, but you can use our course web page and Minitab software.

1. The table below gives some situational OBPs for Rickey Henderson in the 1999 baseball season. This table shows how Rickey did for day and night games, for games home and away, for games played on grass and turf fields, for games played in domed and open ballparks, and against left- and right-handed pitchers. The data are also provided in P:\Data\Math\Hartlaub\SportsStats\RickeyHenderson.mtw.

Breakdown	AB	BB	PA	OBP	Successes
Day	144	21	165	.400	66
Night	294	61	355	.433	153.715
Home	200	34	234	.376	87.98
Away	238	48	286	.462	132.132
Grass	364	60	424	.408	172.992
Turf	74	22	96	.490	47.04
Dome	43	10	53	.472	25.016
Open	395	72	467	.417	194.739
vs. Right	333	54	387	.403	155.961
vs. Left	105	28	133	.481	63.973

a. For each breakdown, compute the number of (approximate) plate appearances (PA). (10)

See Column 4 in the table above.

b. Let p_{day} denote Rickey's on-base probability when he is playing a day game and let p_{night} denote his on-base probability when he's playing at night. Using the appropriate data, construct a 95% confidence interval for the difference between p_{day} and p_{night}. Can you conclude that Rickey really has a higher on-base probability at night games? Explain. (20)

To find the number of successes (times on base) we multiply PA by OBP. Using Stat > Basic Statistics > 2 Proportions $\begin{bmatrix} 1^{st} & 66 & 165 \\ 2^{nd} & 154 & 355 \end{bmatrix}$ the 95% CI for $p_{day} - p_{night}$ is $(-.124607, .0570014)$ Since zero is included in the 95% CI we cannot conclude that Rickey has a higher on base percentage at night.

c. Using the same method as in (b), compare the home/away OBPs, the grass/turf OBPs, the dome/open OBPs, and the right/left OBPs. Be sure to identify the appropriate parameters for each situation. (40 – 10 each situation).

Use Stat > Basic Statistics > 2 Proportions

H/A $\begin{bmatrix} 1^{st} & 88 & 234 \\ 2^{nd} & 132 & 286 \end{bmatrix}$ G/T $\begin{bmatrix} 1^{st} & 173 & 424 \\ 2^{nd} & 47 & 96 \end{bmatrix}$ D/O $\begin{bmatrix} 1^{st} & 25 & 53 \\ 2^{nd} & 195 & 467 \end{bmatrix}$ R/L $\begin{bmatrix} 1^{st} & 156 & 387 \\ 2^{nd} & 64 & 133 \end{bmatrix}$

The parameters for each situation are identified by the subscripts, as in part b. For example, p_{Home} = his on-base probability when he is playing at home.

Home/Away 95% CI for $p_{HOME} - p_{AWAY}$ is $(-.170264, -.000676040)$. Since zero is not in the 95% CI, we have significant evidence that Rickey's OBP is higher when he plays away games.

Grass/Turf 95% CI for $p_{Grass} - p_{Turf}$ is $(-.191963, .028834)$. Since zero is included in the 95% CI we cannot reject the claim that Rickey's OBP is the same on grass and turf fields.

Dome/Open 95% CI for $p_{Dome} - p_{open}$ is $(-.0875032, .195782)$. Since zero is include in the 95% CI, we do not have evidence that Rickey's on-base probability differs in domes and open stadiums.

Right/Left 95% CI for $p_{right} - p_{left}$ is $(-.176076, .0198720)$. Since zero is included in the 95% CI, Rickey's on-base probability is not different for right and left handed pitchers.

2. The data provided in P:\Data\Math\Hartlaub\SportsStats\HomeAway2008.mtw give batting statistics for a sample of 20 players in the 2008 season. The first column gives the name of the player, the next three columns give the number of at-bats, hits, and batting average for all games played at home, and the next three columns give the number of at-bats, hits, and batting average for all games played away from home. Does this sample provide significant evidence that players generally hit better at home? If so, by how much on the average? Provide a complete statistical justification. (30)

Since hitting ability differs from player to player and the same players are examined at home and away, a paired t test or interval must be used. Let diff = Home_Avg − Away_Avg for each player. We are interested in testing $H_0: \mu_d = 0$ vs $H_1: \mu_d > 0$.

μ_d = true mean difference (home − away)

$t = 1.49$, p-value = .077. Since .077 > .05, we cannot reject H_0 and we conclude that the hitters do not generally do better at home. The estimated mean difference is $\bar{x}_d = .0166$ and a 95% CI for μ_d is $(-.0068, .04)$, which includes zero.