Answers to Problem Set

1) You are interested in testing the distance of two golf balls, Brand A and Brand B. You take a random sample of 100 golfers (*sample*), each of whom hits Brand A once and Brand B once. Define X as the distance for Brand A, and define Y as the distance for Brand B, with the distances measured in yards. In your sample of 100 golfers, the sum of X equals 20,000 yards (*statistic*), and the sum of Y equals 19,500 yards (*statistic*). The sum of X^2 equals 4,090,000 yards squared (*statistic*), and the sum of Y^2 equals 3,892,500 yards squared (*statistic*). The sum of the product between X and Y equals 3,963,000 yards squared (*statistic*).

   a. Using the asymptotic distribution, create a 95-percent confidence interval for the average distance of Brand A (*parameter*).
   b. Using the asymptotic distribution, create a 95-percent confidence interval for the average distance of Brand B (*parameter*).
   c. Using the asymptotic distribution, test the hypothesis that the average distance of Brand A equals the average distance of Brand B (*parameter*). Use a two-tailed test and a five-percent level of significance.

a. The sample mean of X equals 20,000/100 or 200. The sample variance equals $4{,}090{,}000/100-200^2$ or 900. Therefore, the 95-percent confidence interval using the asymptotic distribution equals $200 \pm 1.96\times S/\sqrt{n}$ or {194.12,205.88}.
b. The sample mean of Y equals 19,500/100 or 195. The sample variance equals $3{,}892{,}500/100-195^2$ or 900. Therefore, the 95-percent confidence interval using the asymptotic distribution equals $195 \pm 1.96\times S/\sqrt{n}$ or {189.12,200.88}.

Define μ_X as the mean of X, and define μ_Y as the mean of Y. The null hypothesis that $\mu_X=\mu_Y$ can be written equivalently as $\mu_X-\mu_Y$ equals zero, against the alternative that $\mu_X-\mu_Y$ is not equal to zero. Use the difference in the sample means of X and Y as the estimate for $\mu_X-\mu_Y$. An equivalent statistic is the sample mean of W, where w_i equals x_i-y_i for the 100 golfers. Thus, the null hypothesis can be written equivalently as μ_W equals zero, where μ_W is the population mean of W. By the theorem of expectations for linear functions, E(W) equals E(X)–E(Y) and V(W) equals V(X)+V(Y)–2C(X,Y).

c. The sample covariance between X and Y equals $3{,}963{,}000/100-200\times195$ or 630. Using the sample variances for X and Y from above, the sample variance for W equals $900+900-2\times630$ or 540. Because the difference in the sample means of X and Y equals 5, the test statistic for the null hypothesis equals $5/\sqrt{(540/100)}$, which equals 2.1517. The test statistics exceeds 1.96, so reject the null hypothesis.

2) You now structure your sample differently. You take a random sample of 100 golfers (*sample*), each of whom hits Brand A once. You then take a second random sample of 100 golfers (*sample*), each of whom hits Brand B once. The second sample is independent of the first. Continue to define X as the distance for Brand A and Y as the distance for Brand B. Assume the sums and sums of squares remain the same as above, so the sum of X equals 20,000 yards (*statistic*), and the sum of Y equals 19,500 yards (*statistic*), the sum of X^2 equals 4,090,000 yards squared (*statistic*), and the sum of Y^2 equals 3,892,500 yards squared (*statistic*). Under the new structure of the sample, how will the answers to a, b, and c change?

The answers to A and B remain the same with the independent samples, but the answer to C changes. Because the covariance is zero, the variance of the sample mean of X minus the sample mean of Y equals the sum of their variances. Therefore, the test statistic becomes $5/\sqrt{[(900+900)/100]}$, which equals 1.1785. The test statistic is less than 1.96, so you cannot reject the null hypothesis.

Using the same set of golfers introduces correlation between the sample mean for Brand A and the sample mean for Brand B. When the correlation is positive, it reduces the variance in the estimate for the average difference. This increases the power of the hypothesis test and, in this case, allows rejection of the null.