

# Max Greene Prob & Stat HW12

## 9.2.2

```
x_bar <- -4.7; Sx <- 7.05; Sx2 <- Sx^2
y_bar <- -1.6; Sy <- 5.63; Sy2 <- Sy^2
W <- (x_bar-y_bar-0)/(sqrt(Sx2/27+Sy2/27))
theta_hat <- Sx2/Sy2
df <- round((theta_hat+1)^2/((1/(27-1))*(theta_hat^2+1)))
W
```

```
## [1] -1.785389
```

```
qt(c(0.05,0.95),df=df)
```

```
## [1] -1.675905 1.675905
```

Yes, they are significantly different.  $W$  is more extreme than the  $\alpha = 0.05$  confidence interval in the student  $t$  distribution with  $df = 50$

## 9.2.12

```
x_bar <- 81.6; Sx2 <- 17.6
y_bar <- 79.9; Sy2 <- 22.9
W <- (x_bar-y_bar-0)/(sqrt(Sx2/10+Sy2/20))
#theta_hat <- Sx2/Sy2
#df<-round((theta_hat+1)^2/((theta_hat^2/(10-1))+(20/(20*19)^2)))
W
```

```
## [1] 0.9974149
```

## 9.2.14

Prove that the  $Z$  ratio given in Equation 9.2.1 has a standard normal distribution.

$$Z = \frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}}$$

We know that  $\bar{X}, \bar{Y}$  are normal by the mean value theorem because their individual values come from a normal distribution. Adding to and dividing by constants does not change the normal nature of the distribution. We must now show that the expected value and variance are that of a standard normal distribution.

$$\begin{aligned} E(Z) &= \int_{-\infty}^{\infty} xZ(x)dx + \int_{-\infty}^{\infty} yZ(y)dy \\ &= \int_{-\infty}^{\infty} x \frac{\bar{X} - \mu_X}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}} dx + \int_{-\infty}^{\infty} y \frac{\bar{Y} - \mu_Y}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}} dy = 0 \end{aligned}$$

Since  $E(\bar{X} - \mu_X), E(\bar{Y} - \mu_Y) = 0$ .

$$Var(Z) = E(Z^2) - E(Z)^2 = E(Z^2)$$

## 9.2.18

The  $n + m - 2$  degrees of freedom will provide the greater accuracy in this case because of the higher degrees of freedom. Therefore, it will be advantageous and provide a more reliable test to use the  $n + m - 2$  degrees of freedom.

## 9.3.4

By theorem 9.3.1, the test  $H_0 : \sigma_X^2 = \sigma_Y^2$  vs  $H_1 : \sigma_X^2 \neq \sigma_Y^2$  will reject  $H_0$  if  $\frac{s_Y^2}{s_X^2}$  is either  $\leq F_{\alpha/2, m-1, n-1}$  or  $\geq F_{1-\alpha/2, m-1, n-1}$

```
Sx2 <- 5.67^2; Sy2 <- 3.18^2;
data <- data.frame(cage=(1:20), WG=c(22.8, 10.2, 20.8, 27.0, 19.2, 9.0, 14.2, 19.8, 14.5, 14.8, 23.5, 31.0, 19.5, 26.1, 19.2, 19.8, 14.5, 14.8, 23.5, 31.0),
  Sy2/Sx2
```

```
## [1] 0.3145489
```

```
qf(c(0.05/2, 1-0.05/2), 9, 9)
```

```
## [1] 0.2483859 4.0259942
```

No, the variances are not significantly different.

## 9.3.6

If the variables are independent and normally distribution with the same standard deviation then the two-sample  $t$  test would be the correct test to use for this data.

## 9.3.8

Theorem 9.2.2 is best used for two independent random samples from normal distributions with the same standard deviations to test for differences in means. Theorem 9.2.3, on the other hand, is best used for two independent random samples from normal distributions with possibly differing means and differing standard deviations to test for differences in means.

The difference between this test is that the samples in 9.2.3 can have separate sd's and the theorem gives a way to handle this case.

Given that the standard deviations of these populations do not seem to be associated in any way and could be completely independent, theorem 9.2.3 is more appropriate for deciding whether the difference between means is significant.

## 9.4.2

```
x <- 66; y <- 93; n <- 846/2; m <- 846/2;
pe <- (x+y)/(n+m)
(z <- (x/n - y/m)/sqrt((pe*(1-pe))/n + (pe*(1-pe))/m))
```

```
## [1] -2.37614
```

```
qnorm(0.05)
```

```
## [1] -1.644854
```

The diet significantly decreased the chances of arteriosclerosis since  $z \leq -z_\alpha$ .

## 9.4.6

```
n <- 100;m <- 100;x <- 60;y <- 48;
pe <- (x+y)/(n+m)
(z <- (x/n-y/m)/sqrt((pe*(1-pe))/n+(pe*(1-pe))/m))
```

```
## [1] 1.702513
```

```
pnorm(z)
```

```
## [1] 0.9556704
```

the  $p$  value associated with this data is 0.9556 or 95.56%

## 9.5.2

```
dome <- c(100,58.6,93.5,83.6,84.1)
noDome <- c(76.4,84.2,96.5,88.8,79.1,83.6)
sp <- 11.2; x_bar <- mean(dome); y_bar <- mean(noDome);
n <- length(dome); m <- length(noDome); alpha <- 0.05;
temp <- c(qt(alpha/2,n+m-2)*sp*sqrt(1/n+1/m),-qt(alpha/2,n+m-2)*sp*sqrt(1/n+1/m))
x_bar-y_bar+temp
```

```
## [1] -16.14847 14.53514
```

## 9.5.6

95% confidence interval for  $\sigma_X^2/\sigma_Y^2$  based on case study in 9.2.1.

```
alpha <- 0.05
s2x <- .0002103; s2y <- .0000955; ratio <- s2x/s2y;
ratio*qf(c(alpha/2,1-alpha/2),m-1,n-1);
```

```
## [1] 0.2980683 20.6214473
```

## 9.5.10

Construct an 80% confidence interval for difference  $p_M - p_W$  in the nightmare frequency data summarized in Case Study 9.4.2.

```
x <- 55;y <- 60;n <- 160;m <- 192;
zsqr <- qnorm(c(0.1,0.9))*sqrt((x/n)*(1-x/n)/n+(y/m)*(1-y/m)/m)
x/n-y/m+zsqr
```

```
## [1] -0.03319675 0.09569675
```

## 13.3.2

Use a paired  $t$  test with  $\alpha = 0.05$  to compare numbers of trials needed to learn depth perception for Mothered and Unmothered lambs in Question 8.2.6.

```
x <- c(2,3,5,3,2,1,1,5,3,1,7,3,5)
y <- c(3,11,10,5,5,4,2,7,5,4,8,12,7)
d <- x-y
```

```
meanD <- mean(d); sdD <- sd(d); n <- length(x);  
(t <- meanD/(sdD/sqrt(n)))
```

```
## [1] -4.502873
```

```
qt(c(0.05,0.95),df=n-1)
```

```
## [1] -1.782288 1.782288
```

```
pt(t,df=n-1)
```

```
## [1] 0.000361514
```

These samples,  $x$  and  $y$ , are significantly different since their  $t$  value is larger than the upper bound of the 90% confidence interval. Specifically, the differences were in the 99.96th percentile of significance given the null hypothesis  $\mu_X = \mu_Y$  was true.

### 13.3.6

```
d <- c(0.8,1.9,-0.4,1.2,0.2,-0.4,-0.6,0.4,0.4,1.2)  
meanD <- mean(d); sdD <- sd(d); n <- length(d);  
(t <- meanD/(sdD/sqrt(n)))
```

```
## [1] 1.826247
```

```
qt(c(0.025,0.975),df=n-1)
```

```
## [1] -2.262157 2.262157
```

```
pt(t,df=n-1)
```

```
## [1] 0.9494519
```

This difference is not significant with  $\alpha = 0.025$ . There is a (100-94.95)% chance that data this extreme would occur if the null hypothesis, that the two population means were equal, were to be true.