Part 1

The first step is to put the part 1 data into R. One way to do that is to create two vectors containing the test result frequencies for the pregnant and the non-pregnant users. Note that the “yes” responses are the first number and the “no” responses are the second number in each vector.

```r
# frequency vectors
freq.preg <- c(451, 36)  # hits, misses
freq.notpreg <- c(15, 183)  # false alarms, correct rejections
```

Then convert the frequencies to probabilities by dividing each vector by the total number of pregnant women or the total number of not-pregnant women

```r
prob.preg <- freq.preg / sum(freq.preg)
prob.notpreg <- freq.notpreg / sum(freq.notpreg)
```

You can examine the probabilities you just computed by typing the names of the objects (i.e., prob.preg and prob.notpreg. In the example, I have used the round() command to show only 4 decimal places.

```r
round(prob.preg, 4)
```

```
## [1] 0.9261 0.0739
```

```r
round(prob.notpreg, 4)
```

```
## [1] 0.0758 0.9242
```

We see that the E.P.T. had a hit rate of 0.9261 and a false alarm rate of 0.0758. Finally convert the probabilities to z-scores.

```r
zprob.preg <- qnorm(prob.preg)
zprob.notpreg <- qnorm(prob.notpreg)
```

Now we can compute the detection (d-prime), sensitivity (Az) and bias (c) indices using the equations in the Detection Theory handout. The [1] notation after the vector names indicates that we want to use just the first item in each vector. These items correspond to the z-score of the hit rate and the z-score of the false alarm rate.

```r
# d-prime and c
cbias <- -(zprob.preg[1] + zprob.notpreg[1]) / 2  # Equation 12
```
The d’ for discriminating between pregnant (signal) and not pregnant (noise) for this detector is 2.881. The decision bias of the operator evaluating the detector is -0.006, very close to zero. We can use pnorm() function to compute the accuracy, A-sub-z, from the z-score quantiles using Equation 14 in the Detection Theory handout. The function pnorm() converts z-scores back into probabilities.

\[
Az \leftarrow \text{pnorm}\left(\frac{d' \text{ / } \sqrt{2}}{\text{sqrt}(2)}\right) \quad \# \text{ Equation 14}
\]

The results show that the overall accuracy of the E.P.T. is 0.9792. The high accuracy and low bias of the test suggests that it is effective at discriminating between women who are and are not pregnant.

**Part 2**

First, calculate the hit rate for all the detectors, as well as the false alarm rate at different levels of bias, and then make an ROC plot. In the code below the `par(pty = "s")` command forces the plot to be square. The `par(pty = "m")` command after the plot restores graphs to maximum size.

```r
FAR <- c(.84, .62, .34, .21, .05)
HR <- data.frame(S1 = c(.93, .79, .54, .38, .14),
                 S2 = c(.98, .90, .73, .58, .27),
                 S3 = c(.99, .96, .86, .76, .46))

colors <- c("turquoise3","springgreen4","firebrick2")
par(pty = "s")
plot(c(0,1),c(0,1), type="l", xlab="FAR", ylab="HR", main="ROC plot for cancer diagnostics")
points(FAR,HR$S1, col=colors[1], type="b")
points(FAR,HR$S2, col=colors[2], type="b")
points(FAR,HR$S3, col = colors[3], type="b")
legend("bottomright", c("S1","S2","S3"), pch = 1, col=colors)
```
Turn the probabilities into z-scores, and then make another ROC plot of the z-scores.

```r
zHR <- data.frame(S1 = qnorm(HR$S1), S2 = qnorm(HR$S2), S3 = qnorm(HR$S3))
zFAR <- qnorm(FAR)
modS1 <- lm(zHR$S1 ~ zFAR)
modS2 <- lm(zHR$S2 ~ zFAR)
modS3 <- lm(zHR$S3 ~ zFAR)

par(pty = "m")
plot(c(-2.5,2.5),c(-2.5,2.5),
     type="l",
     xlab="z(FAR)",
     ylab="z(HR)",
     main="ROC plot for cancer diagnostics")
points(zFAR,zHR$S1,col=colors[1])
points(zFAR,zHR$S2,col=colors[2])
points(zFAR,zHR$S3,col=colors[3])
abline(a=modS1$coefficients[1],b=modS1$coefficients[2],col=colors[1])
abline(a=modS2$coefficients[1],b=modS2$coefficients[2],col=colors[2])
abline(a=modS3$coefficients[1],b=modS3$coefficients[2],col=colors[3])
legend("bottomright",
       c("S1","S2","S3"),
       pch = 1,
       col=colors)
```
par(pty = "m")

Use pnorm() to calculate the accuracies of the different detectors

AzS1 <- pnorm(as.numeric(modS1$coefficients[1])/sqrt(2))
AzS2 <- pnorm(as.numeric(modS2$coefficients[1])/sqrt(2))
AzS3 <- pnorm(as.numeric(modS3$coefficients[1])/sqrt(2))

The accuracy of S1, S2, and S3 is 0.64, 0.76, and 0.85 respectively. S3 is the best detector of the group.