Math 3070 § 1. **Birth Weight Example:** Name: Example
Treibergs **One Sample z-Test, CI and Power** May 27, 2011

This exercize is about z-tests. For large sample sizes $n > 40$, by rule of thumb, the statistic

$$t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$$

is approximately normal, and hypothesis tests can be made using critical z-values. In fact, for normal data, this statistic satisfies the t-disribution with $n-1$ degrees of freedom, and the t-test, which is canned in **R** is usually used.

We analyze the birth weights observed by Secher in the `ISwR` package. For sake of concreteness, we ask: are the birthweights of European babies significantly above 6 pounds? (I don't know if anyone would ever want to do this!) We calculate the statistic and p-value "by hand" and find that the average is not significantly higher at all. Let $\mu_0 = 2721.6$ grams. This is 6 pounds times 453.6 grams per pound. the null and alternative hypotheses are

$$\begin{aligned} H_0 : &\quad \text{The mean } \mu \le \mu_0; \\ H_a : &\quad \text{The mean } \mu > \mu_0. \end{aligned}$$

For this one-tailed test, we compute the statistic t and reject the null hypothesis if $t > z_\alpha$. At the $\alpha = .05$ significance level, $z_\alpha = \Phi^{-1}(1-\alpha)$ or about $z_{.05} = 1.645$.

We then go on to find β, the probability of a Type II error. In the case of z tests, this can be computed using normal probabilities. If population mean is μ_1 instead, $\beta(\mu_1) = P(H_0 \text{ is accepted} | \mu = \mu_1)$. This becomes

$$\beta = P\left(T \le z_\alpha | \mu = \mu_1\right) = P\left(\frac{\bar{X} - \mu_0}{s/\sqrt{n}} \le z_\alpha | \mu = \mu_1\right) = \Phi\left(z = \frac{\bar{X} - \mu_1}{s/\sqrt{n}} \le z_\alpha - d\right)$$

where the normalized difference is

$$d = \frac{\mu_1 - \mu_0}{s/\sqrt{n}}.$$

The two tailed situation is similar. This time $z_{.025} = 1.960$ and

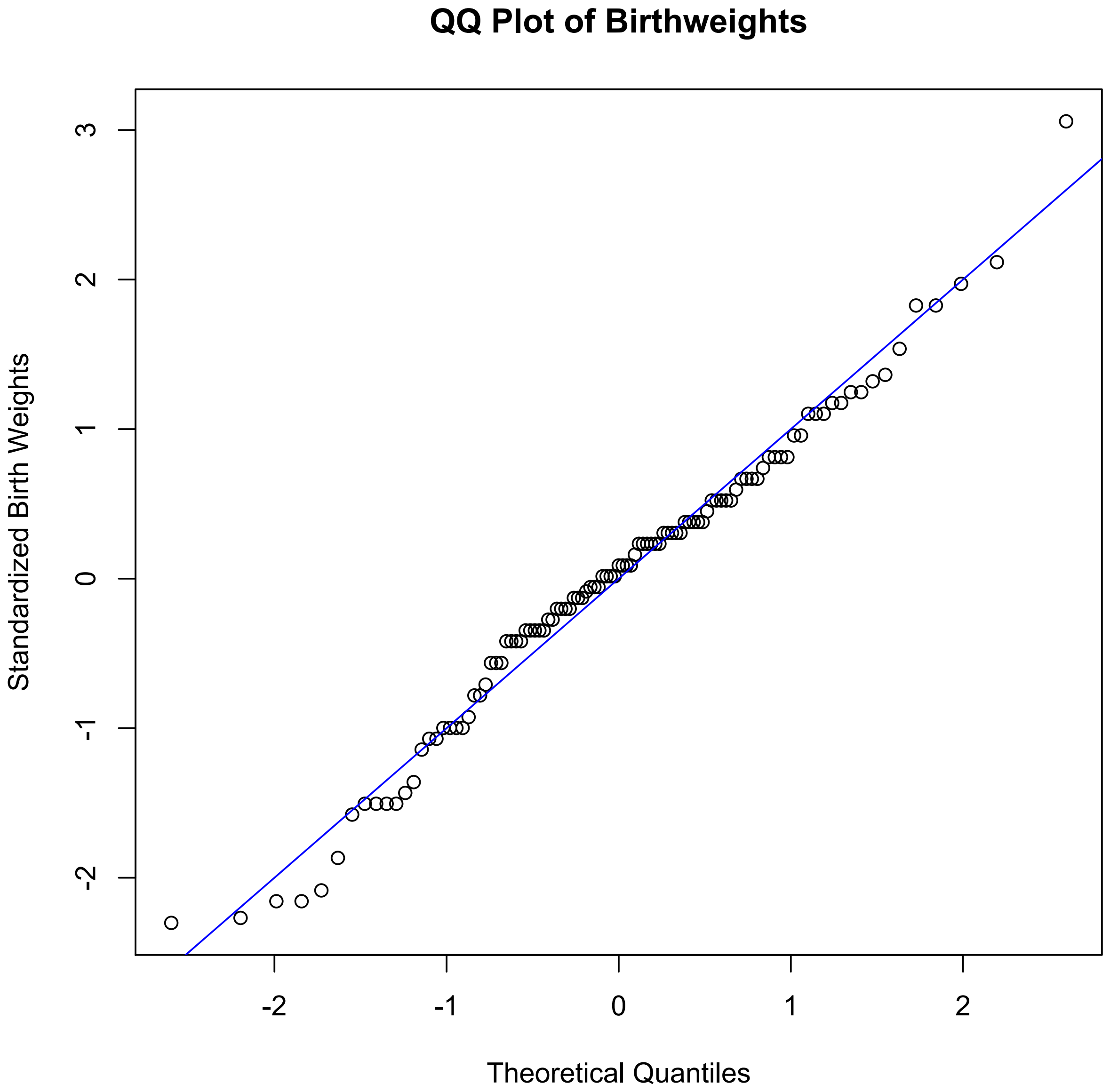
$$\begin{aligned} \beta &= P\left(-z_{\alpha/2} \le T \le z_{\alpha/2} | \mu = \mu_1\right) \\ &= P\left(-z_{\alpha/2} \le \frac{\bar{X} - \mu_0}{s/\sqrt{n}} \le z_{\alpha/2} | \mu = \mu_1\right) \\ &= \Phi\left(z \le z_{\alpha/2} - d\right) - \Phi\left(z \le -z_{\alpha/2} - d\right) \end{aligned}$$

We plot curves showing how β depends on d. We also plot normal curve diagrams of the areas these probabilities correspond to.

R Session:

```
R version 2.11.1 (2010-05-31)
Copyright (C) 2010 The R Foundation for Statistical Computing
ISBN 3-900051-07-0

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.
```

```
  Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[R.app GUI 1.34 (5589) i386-apple-darwin9.8.0]

[Workspace restored from /home/1004/ma/treibergs/.RData]

> # Do random European babies exceed 6 lb birthweight?
> # Secher 1987 study of European babies is a canned dataset
> # secher  in the package ISwR. Load the package and look at
> # the first few lines. The variable  bwt  is the birthweight.
>
> library(ISwR)
> head(secher)
   bwt bpd  ad no
1 2350  88  92  1
2 2450  91  98  2
3 3300  94 110  3
4 1800  84  89  4
5 2900  89  97  5
6 3500 100 110  6
>
> attach(secher)
> summary(bwt)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   1150    2400    2800    2739    3125    4850

> n <- length(bwt);n
[1] 107


> xbar <- mean(bwt)
> s <- sd(bwt)
> # Standardizing means that the QQ points will line up on y = 0 + 1 * x:
> z <- (bwt-xbar)/s
> qqnorm(z,ylab="Standardized Birth Weights",main="QQ Plot of Birthweights")
> abline(0,1,col=4)
> # M3074BirthWeight1.pdf
```

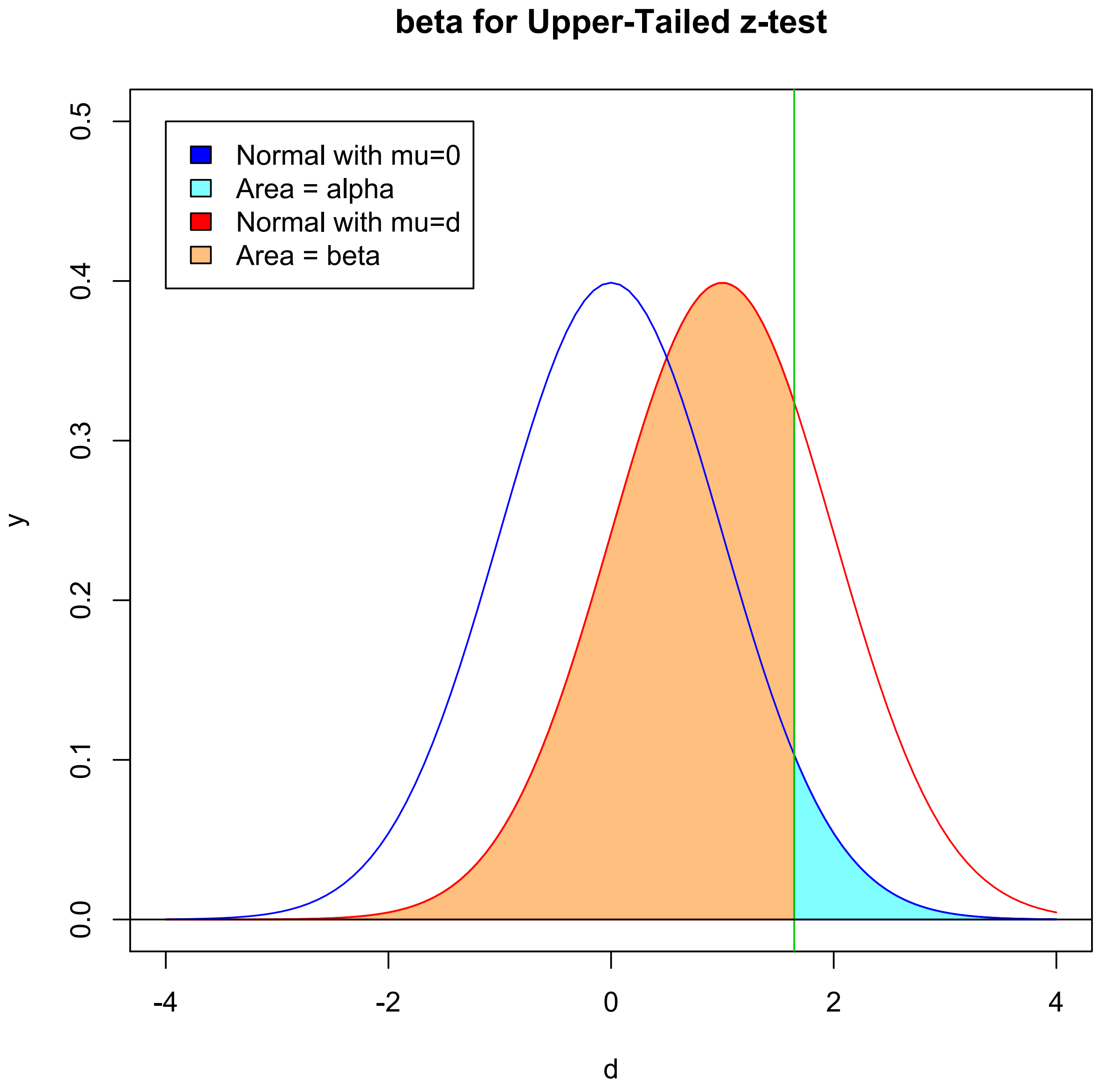


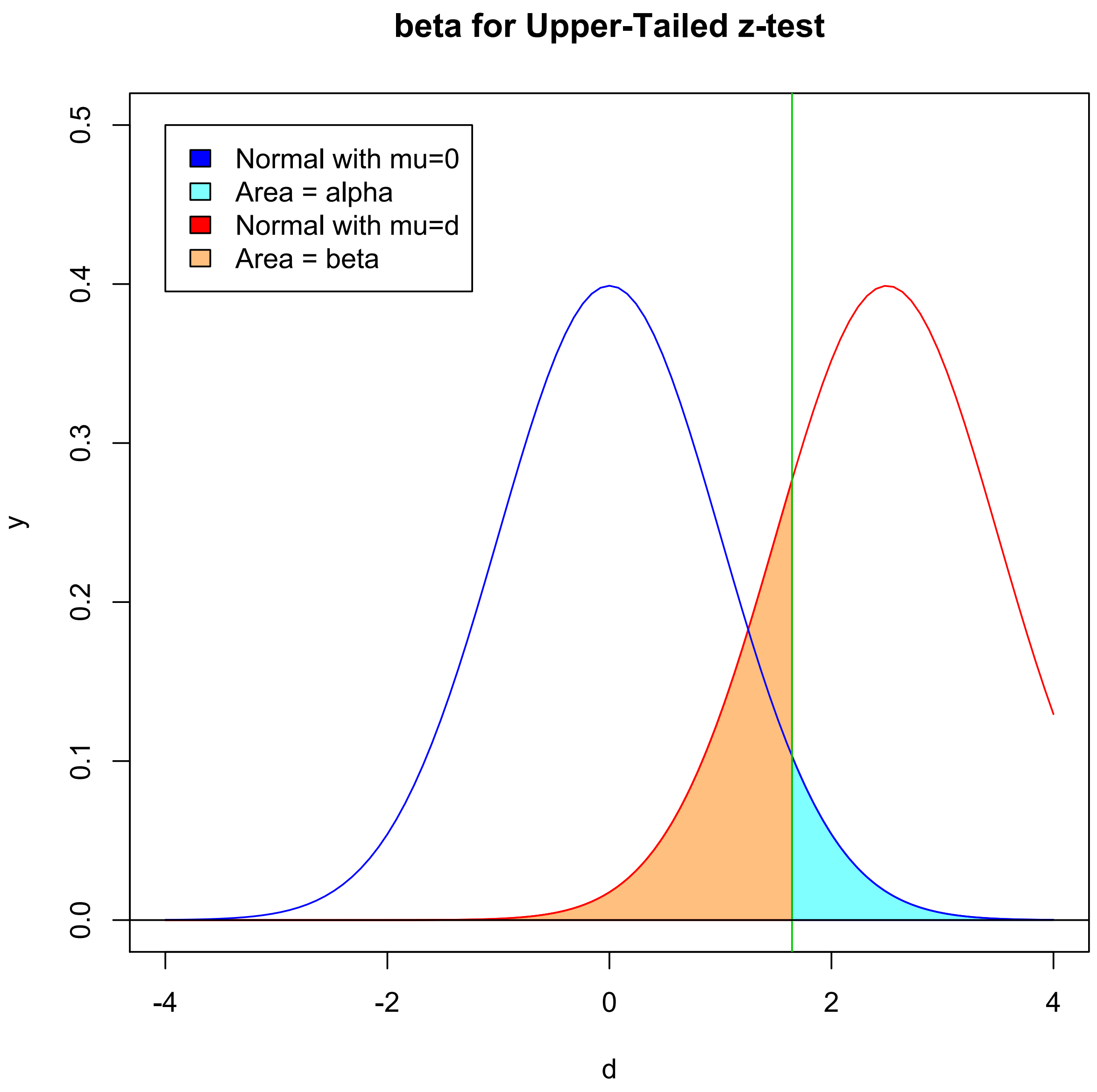
```
> # Looks pretty notmal!
>
> # there are  453.6  grams / pound so 6 lb is in grams
> mu0 <- 6 * 453.6;mu0
[1] 2721.6


> # Do random European babies exceed 6 lb birthweight?
> # Do random European babies significantly exceed 6 lb birthweight?
> # Compute the  T  statistic and the  p-value
> T <- (xbar-mu0)/(s/sqrt(n)); T
[1] 0.2621352
```

```
> pv <- pnorm(T,lower.tail=F); pv
[1] 0.3966086
>
> # So we can’t reject the null hypothesis.
>
> # 95% lower CI for birthweight
> alp <- 0.05; alp
[1] 0.05
> zalpha <- qnorm(alp,lower.tail=F); zalpha
[1] 1.644854

> lci <- xbar - zalpha * s/sqrt(n); lci
[1] 2629.325
```

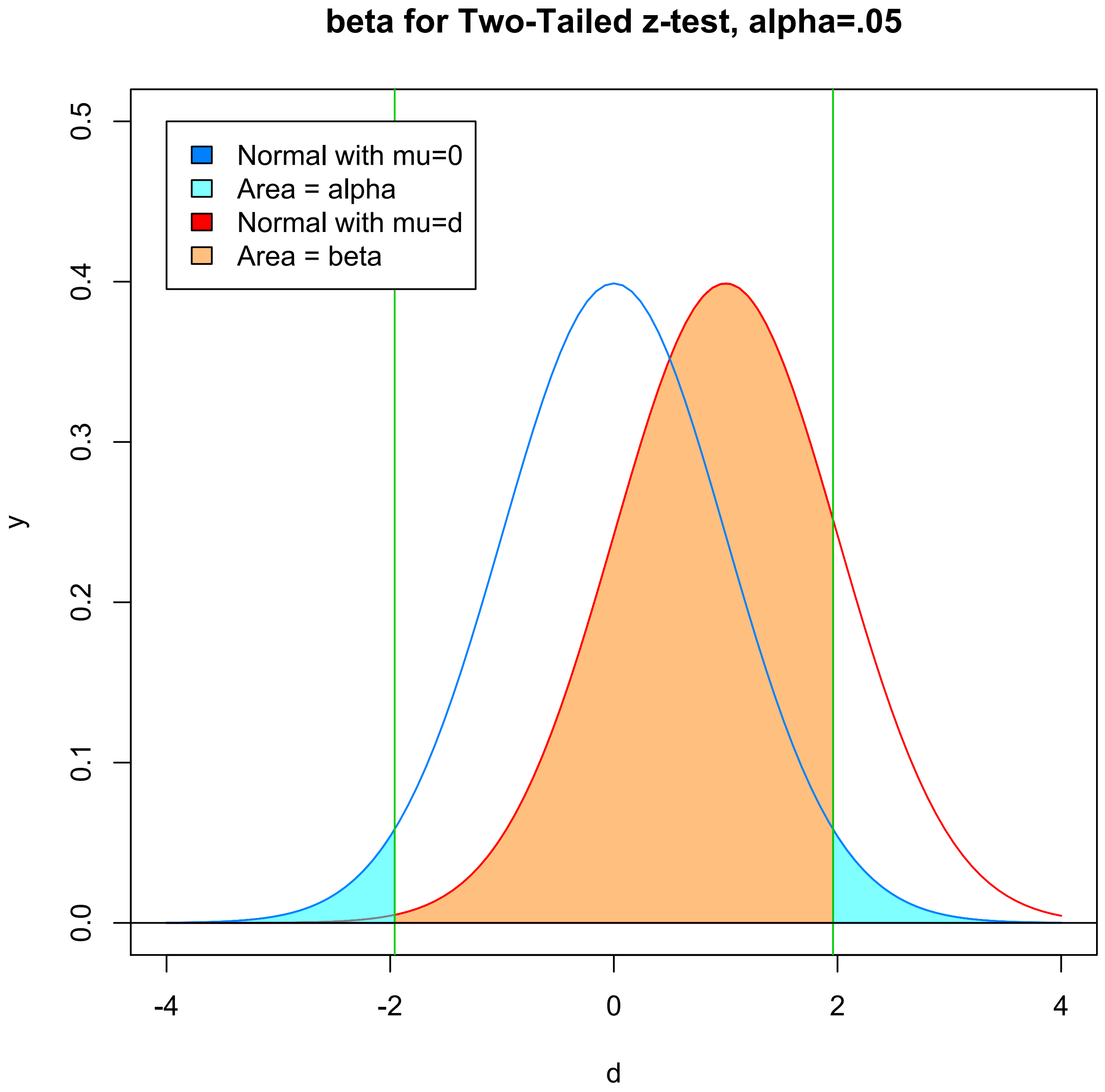
```
> # We can make it look like  R  output.
>
> cat("   One Sample z-test\n\n z = ", T, ", p-value =", pv,
+ "\n alternative hypothesis: true mean is greater than", mu0,
+ "\n Level ",alp,"confidence interval:\n",lci,
+ "   Inf\n sample estimate:\n mean of x\n",xbar,"\n\n")


   One Sample z-test

 z =  0.2621352 , p-value = 0.3966086
 alternative hypothesis: true mean is greater than 2721.6
 Level 0.05 confidence interval:
 2629.325    Inf
 sample estimate:
 mean of x
 2739.093

> # Running the  t-test yields almost the same.
>
> t.test(bwt,alternative="greater",mu = 2721.6)


    One Sample t-test

data:  bwt
t = 0.2621, df = 106, p-value = 0.3969
alternative hypothesis: true mean is greater than 2721.6
95 percent confidence interval:
 2628.357      Inf
sample estimates:
mean of x
 2739.093
```

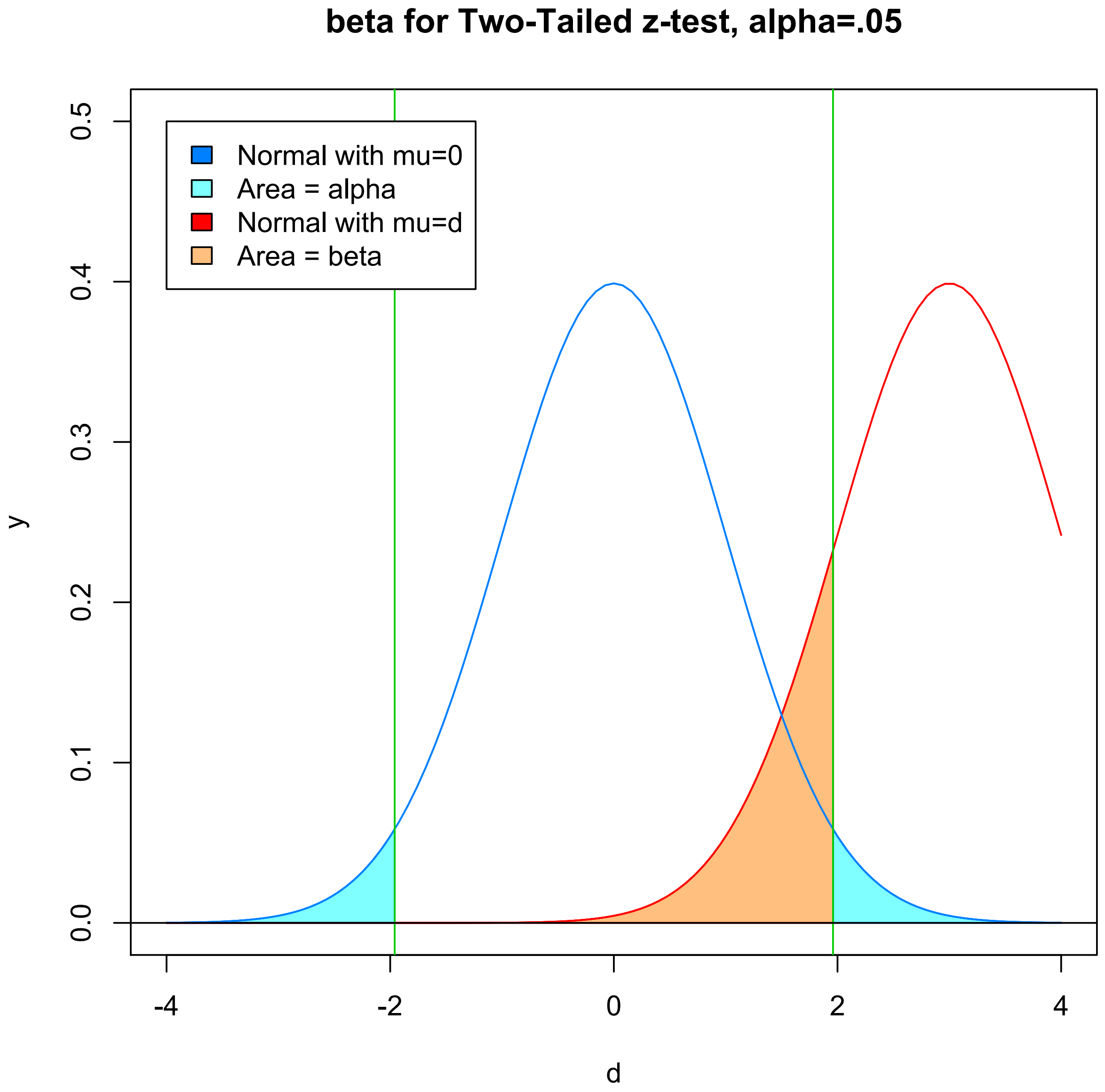
```
> #################### BETA COMPUTATIONS ############################################
>
> # If true mean weight was mu1 then for
> #  d = (mu1-mu0)/(sigma/sqrt(n) )   we have
> # beta = P(z <= zalpha - d)
>
> d =seq(from=.1,to=2, by=.1);d
 [1] 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0 1.1 1.2 1.3 1.4 1.5 1.6 1.7 1.8 1.9 2.0
> pnorm(zalpha-d)
 [1] 0.9388092 0.9257505 0.9106637 0.8934072 0.8738651 0.8519547 0.8276332 0.8009037
 [9] 0.7718199 0.7404890 0.7070729 0.6717872 0.6348978 0.5967151 0.5575868 0.5178880
[17] 0.4780109 0.4383530 0.3993050 0.3612400

> # For example, if the actual mu = mu0 = 3628.8 gm (=8 lb) then the probability of
> # accepting mu <= 6 lb is
> pnorm(zalpha - (3628.8-2721.6)/(s/sqrt(n)))
[1] 3.273038e-33
> # For example, if the actual mu = mu0 = 3175.2 gm (= 7 lb) then the probability of
> # accepting mu <= 6 lb is
> pnorm(zalpha - (3175.2-2721.6)/(s/sqrt(n)))
[1] 1.287023e-07
> # for example, if the actual mu = mu0 = 2948.4 gm (= 6.5 lb) then the probability of
> # accepting mu <= 6 lb is
> pnorm(zalpha - (2948.4-2721.6)/(s/sqrt(n)))
[1] 0.0397419

> ################### NORMAL CURVES SHOWING ALPHA AND BETA ###########################
> # Make a function that inputs a,b,c , fill color and line color, and outputs a normal
> # density curve with mean  c  and sd = 1 and color in the region under the curve from
> #  a  to  b.  In fact, the coloring is done by making a polygon with 100 points along
> # the bell curve.
>
> plo <- function(a,b,c,co,cl){
+                              xx <- c(a,seq(a,b,(b-a)/100),b,a)
+                              yy <- c(0,dnorm(seq(a,b,(b-a)/100)-c),0,0)
+                              curve(dnorm(x-c),-4,4,add=T,col=cl)
+                              polygon(xx,yy,col=co,border=cl)
+                             }
>
> # Now, plot the alpha and beta areas for two different  d's. beta decreases as  d  increases.
> plot(x,y,type="n",xlim=c(-4,4),ylim=c(0,.5),main="beta for Upper-Tailed z-test",xlab="d")
> col1 <- rainbow(12,alpha=.5)[2];coll1 <- rainbow(12)[1]
> col2 <- rainbow(12,alpha=.5)[7];coll2 <- rainbow(12)[9]
> plo(-4,zalpha,1,col1,coll1)
> plo(zalpha,4,0,col2,coll2)
> legend(-4,.5,legend=c("Normal with mu=0","Area = alpha","Normal with mu=d","Area = beta"),
+ fill=c(coll2,col2,coll1,col1))
> abline(v=zalpha,col=3);abline(h=0)
> # M3074BirthWeight2.pdf
>
>
> plot(x,y,type="n",xlim=c(-4,4),ylim=c(0,.5),main="beta for Upper-Tailed z-test",xlab="d")
> plo(-4,zalpha,2.5,col1,coll1)
> plo(zalpha,4,0,col2,coll2)
> legend(-4,.5,legend=c("Normal with mu=0","Area = alpha","Normal with mu=d","Area = beta"),
+ fill=c(coll2,col2,coll1,col1))
> abline(v=zalpha,col=3);abline(h=0)
> # M3074BirthWeight3.pdf
```



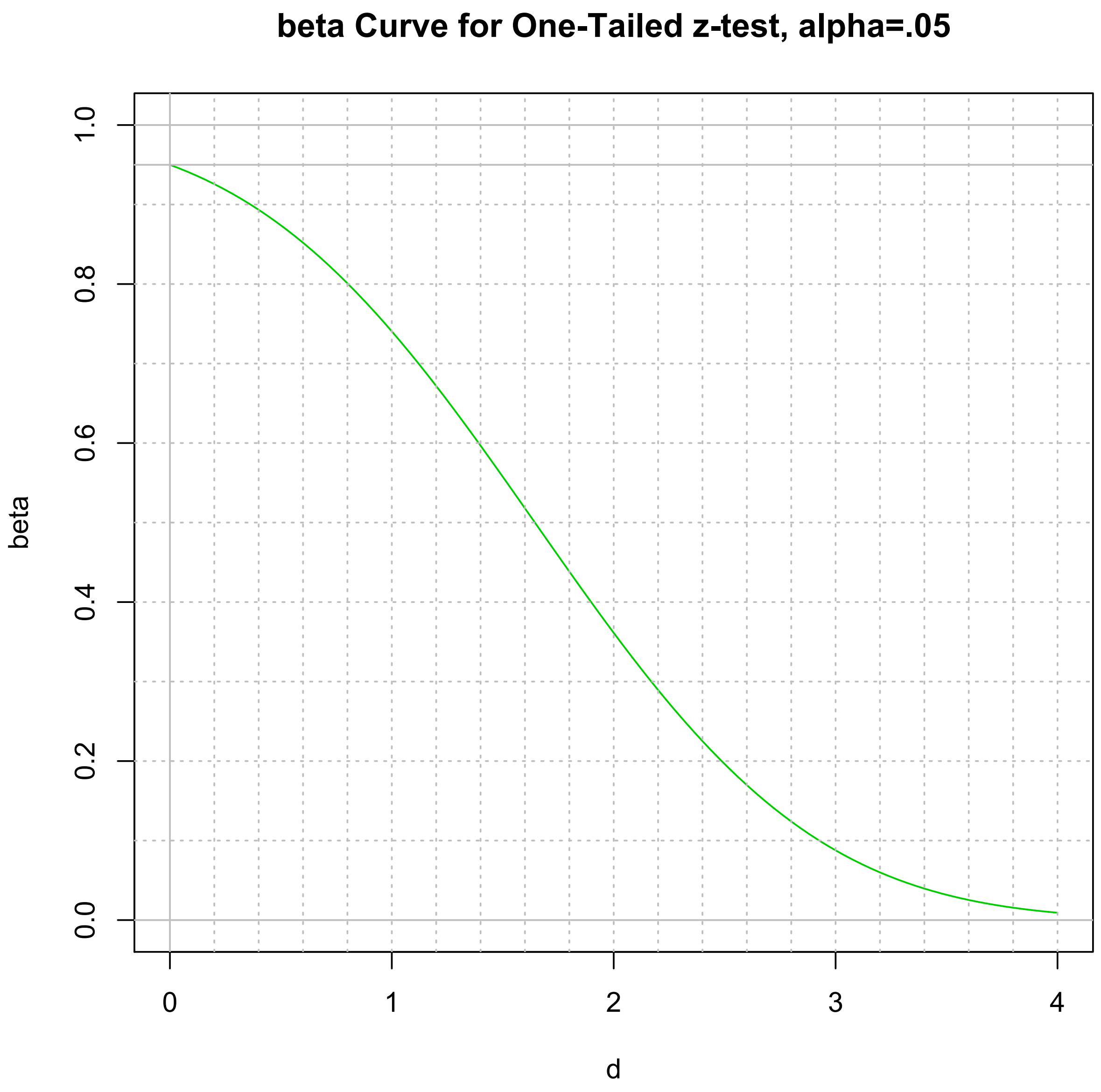


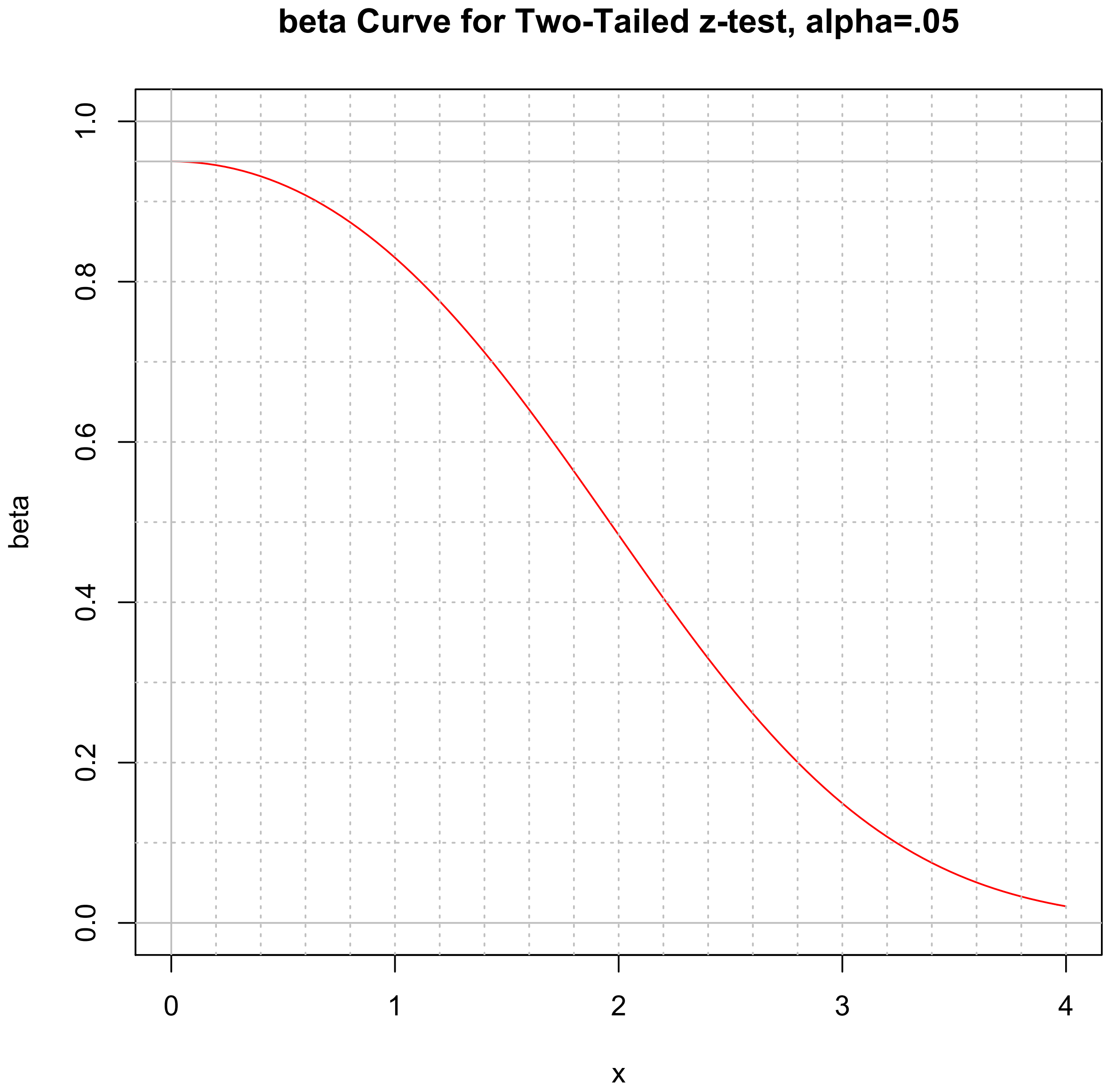
```
> # FOR THE TWO-TAILED TEST WE ALSO HAVE SIMILAR DIAGRAMS.
> za2 <- qnorm(.025,lower.tail=F)
> za2
[1] 1.959964
>
> plot(x,y,type="n",xlim=c(-4,4),ylim=c(0,.5),main="beta for Two-Tailed z-test,
+ alpha=.05",xlab="d")
> plo(-4,za2,1,col1,coll1)
> plo(-4,-za2,1,"white",coll1)
> plo(za2,4,0,col2,coll2)
> plo(-4,-za2,0,col2,coll2)
> abline(v=c(-za2,za2),col=3);abline(h=0)
> legend(-4,.5,legend=c("Normal with mu=0","Area = alpha","Normal with mu=d","Area = beta"),
+ fill=c(coll2,col2,coll1,col1),bg="white")
> # M3074BirthWeight6.pdf
>
> plot(x,y,type="n",xlim=c(-4,4),ylim=c(0,.5),main="beta for Two-Tailed z-test,
+ alpha=.05",xlab="d")
> plo(-4,za2,3,col1,coll1)
> plo(-4,-za2,3,"white",coll1)
> plo(za2,4,0,col2,coll2)
> plo(-4,-za2,0,col2,coll2)
> abline(v=c(-za2,za2),col=3);abline(h=0)
> legend(-4,.5,legend=c("Normal with mu=0","Area = alpha","Normal with mu=d","Area = beta"),
+ fill=c(coll2,col2,coll1,col1),bg="white")
> # M3074BirthWeight7.pdf
```


beta for Two-Tailed z-test, alpha=.05
Normal with mu=0
Area = alpha
Normal with mu=d
Area = beta
y
0.0
0.1
0.2
0.3
0.4
0.5
-4
-2
0
2
4
d


beta for Two-Tailed z-test, alpha=.05
Normal with mu=0
Area = alpha
Normal with mu=d
Area = beta
y
0.0
0.1
0.2
0.3
0.4
0.5
-4
-2
0
2
4
d

```
> ##################### beta AS A FUNCTION OF d #######################################
> # One tailed computation beta computation.
> # If true mean weight was mu1
> # let d = (mu1-mu0)/(sigma/sqrt(n) ) then
> # beta = P(z <= zalpha - d)
>
> curve(pnorm(zalpha-x),0,4,main="beta Curve for One-Tailed z-test, alpha=.05",
+ xlim=c(0,4),ylim=0:1,col=3,ylab="beta",xlab="d")
> abline(v=0,col="gray");abline(v=1:20/5,col="gray",lty=3)
> abline(h=c(0,.95,1),col="gray");abline(h=1:9/10,col="gray",lty=3)
> # M3074BirthWeight4.pdf
>
> # Two tailed computation beta computation.
> # If true mean weight was mu1 then
> # beta = P( -za2 <= X <= za2 | mu = d)  or
> # beta = P(-za2 - d <= z <= za2-d)
>
>  d =seq(from=.1,to=2, by=.1);d
 [1] 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0 1.1 1.2 1.3 1.4 1.5 1.6 1.7 1.8 1.9 2.0
>
> be <- function(x,za){pnorm(za-x)-pnorm(-za-x)}
> be(d,za2)
 [1] 0.9488537 0.9454053 0.9396274 0.9314774 0.9209025 0.9078452 0.8922514 0.8740779
 [9] 0.8533011 0.8299250 0.8039887 0.7755730 0.7448044 0.7118582 0.6769588 0.6403775
[17] 0.6024281 0.5634603 0.5238511 0.4839947
> col1 <- rainbow(12,alpha=.5)[4];coll1 <- rainbow(12)[7]
> col2 <- rainbow(12,alpha=.5)[10];coll2 <- rainbow(12)[11]
> curve(be(x,za2),0,4,main="beta Curve for Two-Tailed z-test, alpha=.05",
+ xlim=c(0,4),ylim=0:1,col=2,ylab="beta")
> abline(v=0,col="gray");abline(v=1:20/5,col="gray",lty=3)
> abline(h=c(0,.95,1),col="gray");abline(h=1:9/10,col="gray",lty=3)
> # M3074BirthWeight5.pdf
```


beta Curve for One-Tailed z-test, alpha=.05
beta
1.0
0.8
0.6
0.4
0.2
0.0
0
1
2
3
4
d


beta Curve for Two-Tailed z-test, alpha=.05
beta
0.0
0.2
0.4
0.6
0.8
1.0
x
0
1
2
3
4