Testing

Oliver d'Pug

Simple vs. Simple

Consider $X_i \overset{iid}{\sim} N(\mu, 2^2)$ with $n = 25$. We wish to test $H_0 : \mu = 0$ versus $H_1 : \mu = 1$. We need to choose k so that

$$\alpha = P\left(\overline{X} > k \mid \mu = 0\right) = 1 - \Phi\left(\frac{k - 0}{2/\sqrt{25}}\right)$$

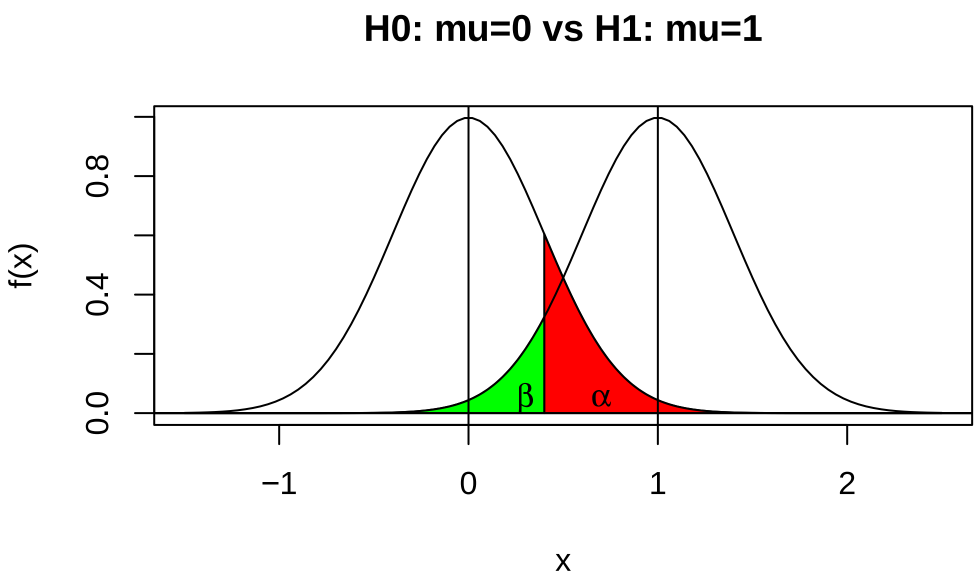
and

$$\beta = P\left(\overline{X} < k \mid \mu = 1\right) = \Phi\left(\frac{k - 1}{2/\sqrt{25}}\right)$$

In pictures, for $k = 0.4$ this looks like:

```
x_max = 2.5
x_min = -1.5
cord.x <- c(.4,seq(.4,x_max,0.01),3)
cord.y <- c(0,dnorm(seq(.4,x_max,0.01),0,2/5),0)
curve(dnorm(x,0,2/5),xlim=c(x_min, x_max),
      main='H0: mu=0 vs H1: mu=1', ylab="f(x)", xlab="x")
polygon(cord.x,cord.y,col='red')
cord.x <- c(-2,seq(x_min,.4,0.01),.4)
cord.y <- c(0,dnorm(seq(x_min,.4,0.01),1,2/5),0)
curve(dnorm(x,1,2/5),xlim=c(x_min, x_max),
      add=TRUE)
polygon(cord.x,cord.y,col='green')

abline(v=c(0,1), lty=1)
abline(h=0, lty=1)
text(0.3,0.05,expression(beta))
text(0.7,0.05,expression(alpha))
```




```
#lines(c(0.6,0.6),c(0,dnorm(0.6,1,2/5)),lty=3)
```

For $k = .4$ we have

```
(sigma = 2)
```

```
[1] 2
```

```
(n = 25)
```

```
[1] 25
```

```
(se = sigma/sqrt(n))
```

```
[1] 0.4
```

```
(k = 0.4)
```

```
[1] 0.4
```

```
(mu = 0) ### H_0 true
```

```
[1] 0
```

```
(alpha = pnorm(k, mu, se, lower.tail=FALSE))
```

```
[1] 0.1586553
```

```
(mu = 1) ### H_1 true
```

```
[1] 1
```

```
(beta = pnorm(k, mu, se, lower.tail=TRUE))
```

```
[1] 0.0668072
```

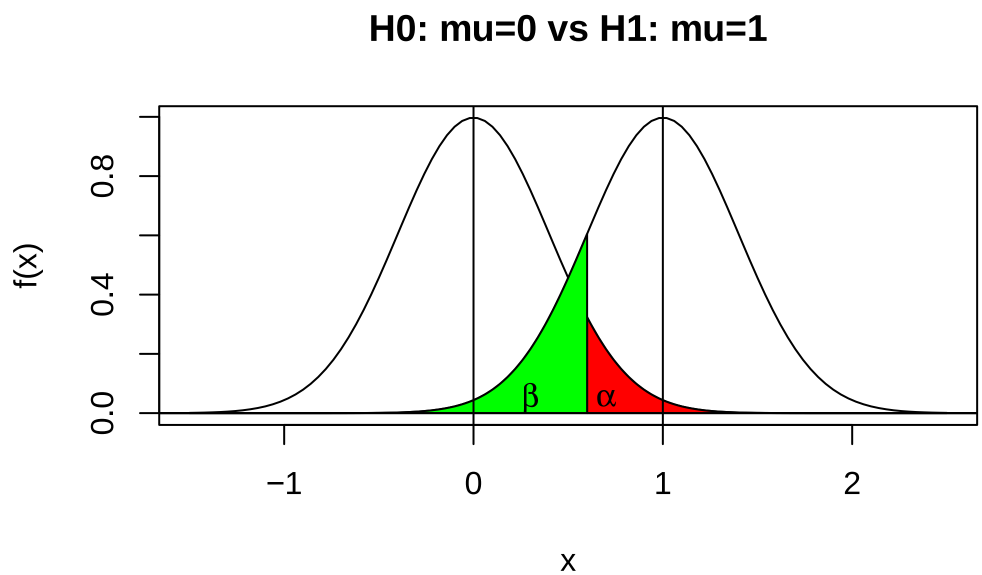
```
(power = 1 - beta)
```

```
[1] 0.9331928
```

For $k = .6$ we have:

```
x_max = 2.5
x_min = -1.5
k = 0.6
cord.x <- c(k,seq(k,x_max,0.01),3)
cord.y <- c(0,dnorm(seq(k,x_max,0.01),0,2/5),0)
curve(dnorm(x,0,2/5),xlim=c(x_min, x_max),
      main='H0: mu=0 vs H1: mu=1', ylab="f(x)", xlab="x")
polygon(cord.x,cord.y,col='red')
cord.x <- c(-2,seq(x_min,k,0.01),k)
cord.y <- c(0,dnorm(seq(x_min,k,0.01),1,2/5),0)
curve(dnorm(x,1,2/5),xlim=c(x_min, x_max),
      add=TRUE)
polygon(cord.x,cord.y,col='green')

abline(v=c(0,1), lty=1)
abline(h=0, lty=1)
text(0.3,0.05,expression(beta))
text(0.7,0.05,expression(alpha))
```




```
#lines(c(0.6,0.6),c(0,dnorm(0.6,1,2/5)),lty=3)
```

```
(k = 0.6)
```

```
[1] 0.6
```

```
(mu = 0) ### H_0 true
```

```
[1] 0
```

```
(alpha = pnorm(k, mu, se, lower.tail=FALSE))
```

```
[1] 0.0668072
```

```
(mu = 1) ### H_1 true
```

```
[1] 1
```

```
(beta = pnorm(k, mu, se, lower.tail=TRUE))
```

```
[1] 0.1586553
```

```r
(power = 1 - beta)
```

```
[1] 0.8413447
```

Coin "Thump" Example

Suppose that we "thump" a "fair" coin ten times. For testing $H_0 : p = 0.5$ versus the one-sided alternative $H_1 : p > 0.5$ at the $\alpha = 0.05$ level we note that $P(X > 7 \mid p = 0.5) = 1 - P(X \leq 7 \mid p = 0.5) =$

```r
1 - pbinom(7, 10, 0.5)
```

```
[1] 0.0546875
```

```r
pbinom(7, 10, 0.5, lower.tail = FALSE)
```

```
[1] 0.0546875
```

Since $0.0546875 > \alpha = 0.05$ we check $P(X > 8 \mid p = 0.5) = 1 - P(X \leq 7 \mid p = 0.5) =$

```r
1 - pbinom(8, 10, 0.5)
```

```
[1] 0.01074219
```

```r
pbinom(8, 10, 0.5, lower.tail = FALSE)
```

```
[1] 0.01074219
```

Since $0.0107422 < \alpha = 0.05$ we reject H_0 when $X > 8$, "flip a coin" with $P(reject) = p^*$ when $X = 8$, and do not reject H_0 when $X \leq 7$.

To find p^* we note that [

$$\begin{aligned} P(rej\ H_0 \mid p = 0.5) &= 0 \cdot P(X \leq 7) + 0.8952 \cdot P(X = 8) + 1 \cdot P(X > 8) \\ &= 0 + 0.0393 + 0.017 \\ &= 0.05 \end{aligned}$$

]

So randomizer coin has probability $p^* = 0.8952$ of causing rejection.

Power

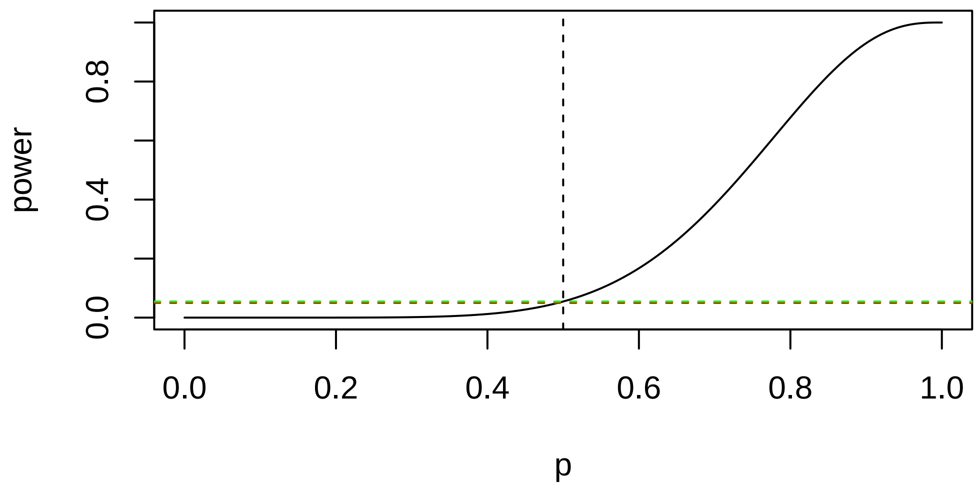
Suppose that we choose $\zeta = \{8, 9, 10\}$ — without randomization at 8. Then $\alpha = P(rej\ H_0 \mid p = 0.5) = 0.0547$

```
pstar <- seq(0, 1, by=0.1)
power <- 1 - pbinom(7, 10, pstar)
print(data.frame(pstar, power = round(power,4)))
```

```
   pstar  power
1    0.0 0.0000
2    0.1 0.0000
3    0.2 0.0001
4    0.3 0.0016
5    0.4 0.0123
6    0.5 0.0547
7    0.6 0.1673
8    0.7 0.3828
9    0.8 0.6778
10   0.9 0.9298
11   1.0 1.0000
```

We can plot the power curve for various "true" p when testing $H_0 : p \leq p_0 = 0.5$ versus $H_1 : p > p_0 = 0.5$.

```
p <- seq(0,1,by=0.001)
power <- 1 - pbinom(7, 10, p)
plot(p, power, type="l")
abline(v=0.5, lty=2)
abline(h=0.05, lty=2, col="red") ### Desired alpha
abline(h=1-pbinom(7,10,0.5), lty=2, col="green")  ### True alpha
```




```
(1-pbinom(7, 10, 0.5)) - 0.05 ### Difference between true and desired alpha
```

```
[1] 0.0046875
```

Note that as p increases (moves further away from ω and into $\Theta - \omega$) our power $(1 - \beta)$ increases. This is a good thing.