Lab 006 - Power Calculations - Solutions

EPIB607 - Inferential Statistics[a]

[a]McGill University

This version was compiled on October 28, 2021

R Code	Value
`qnorm(p = c(0.05, 0.95))`	-1.64, 1.64
`qnorm(p = c(0.025, 0.975))`	-1.96, 1.96
`qnorm(p = c(0.005, 0.995))`	-2.58, 2.58
`qt(p = c(0.025, 0.975), df = 400-1)`	-1.97, 1.97
`qt(p = c(0.025, 0.975), df = 25-1)`	-2.06, 2.06
`qt(p = c(0.025, 0.975), df = 20-1)`	-2.09, 2.09
`qt(p = c(0.025, 0.975), df = 16-1)`	-2.13, 2.13

1. Lake Wobegon

It is claimed that the children of Lake Wobegon are above average. Take a simple random sample of 9 children from Lake Wobegon, and measure their IQ to obtain a sample mean of 112.8. IQ scores are scaled to be Normally distributed with mean 100 and standard deviation 15.

a) Does this sample provide evidence to reject the null hypothesis of no difference between children of Lake Wobegon and the general population?

$$H_0 : \mu = 100 \qquad H_A : \mu > 100$$

The p-value (one-sided test) is given by:

```
pnorm(q = 112.8, mean = 100, sd = 15 / sqrt(9), lower.tail = FALSE)
```

```
#  [1] 0.005233608
```

This sample provides evidence against the null hypothesis. We calculated a p-value of 0.01 which tells us the probability of observing the sample mean of 112.8 under the null hypothesis distribution is very unlikely.

b) Suppose you hope to use a one-sided test to show that the children from Lake Wobegon are at least 10 points higher than average on the IQ test. What power do you have to detect this with the sample of 9 children if using a 0.05-level test?

$$H_0 : \mu = 100 \qquad H_A : \mu > 110$$

Step 1 is to calculate the cutoff in order to reject the null. This is given by

```r
# this is a one-sided alternative so we want alpha=5% in the right tail
(cutoff <- qnorm(p = 0.95, mean = 100, sd = 15 / sqrt(9)))
```

```
#  [1] 108.2243
```

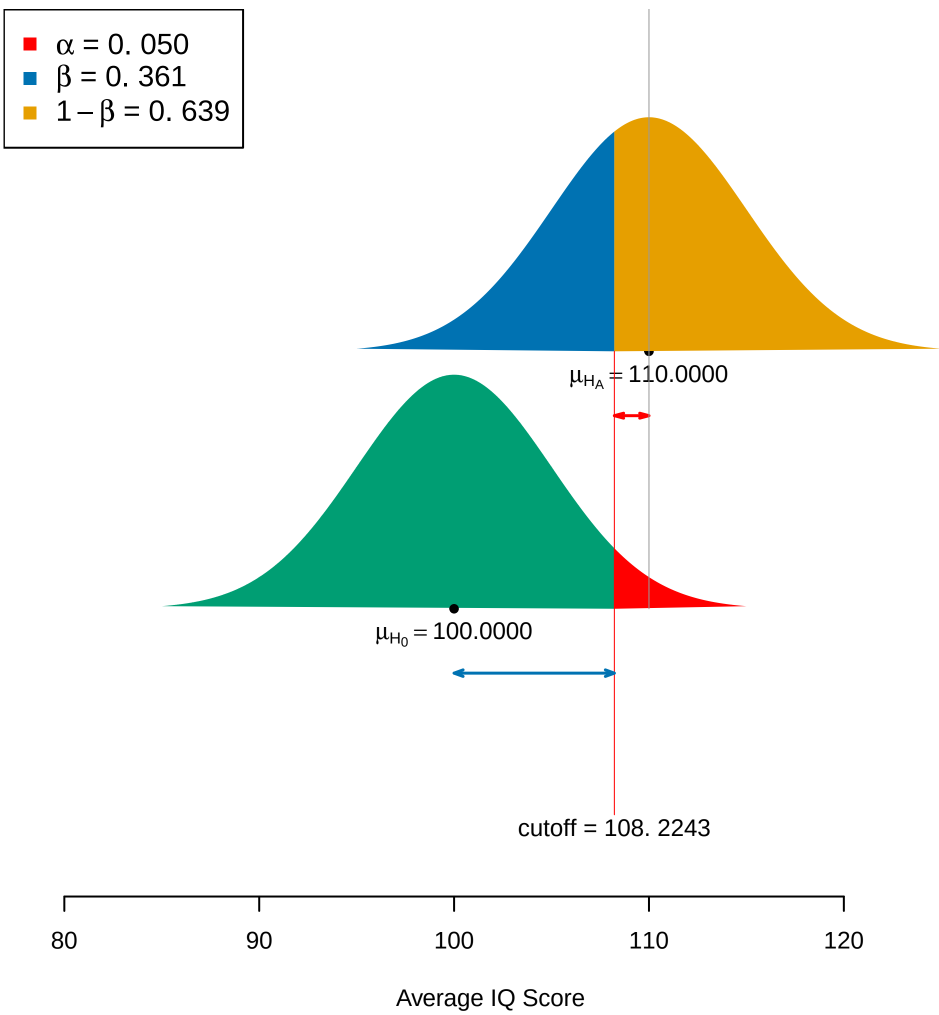
Then we calculate the probability of observing this cutoff or greater under the alternative hypothesis:

```r
pnorm(q = cutoff, mean = 110, sd = 15 / sqrt(9), lower.tail = FALSE)
```

```
#  [1] 0.63876
```

The following figure visualizes this calculation:

```r
source("https://raw.githubusercontent.com/sahirbhatnagar/EPIB607/master/inst/code/plot_null_alt.R
power_plot(n = 9, s = 15, mu0 = 100, mha = 110,
           cutoff = qnorm(p = 0.95, mean = 100, sd = 15 / sqrt(9)),
           alternative = "greater", xlab = "Average IQ Score")
```



c) If you hoped to use a **two-sided** test to show that the children from Lake Wobegon are at least 5 points higher than average on the IQ test, what power do you have with the sample size of 9 and a 0.05-level test?

$$H_0 : \mu = 100 \qquad H_A : \mu = 105$$

Because its a two-sided test, we need to find the both cutoffs, i.e., the values of the sample mean that will reject the null. This is given by:

```
# two-sided test at alpha=5% means we want 2.5% in the tails
(cutoffs <- qnorm(c(0.025, 0.975), 100, 15 / sqrt(9)))
```

```
#  [1]  90.20018 109.79982
```

That is, we will reject the null hypothesis if the sample mean is 90.2 or less, OR reject than null if the sample mean is 109.8 or more.

Next we need to calculate these probabilities under the alternative hypothesis:

```
# left tail probability
(p_left <- pnorm(q = 90.20018, mean = 105, sd = 15 / sqrt(9), lower.tail = TRUE))
```

```
#  [1] 0.001538375
```

```
# right tail probability
(p_right <- pnorm(q = 109.79982, mean = 105, sd = 15 / sqrt(9), lower.tail = FALSE))
```

```
#  [1] 0.1685367
```

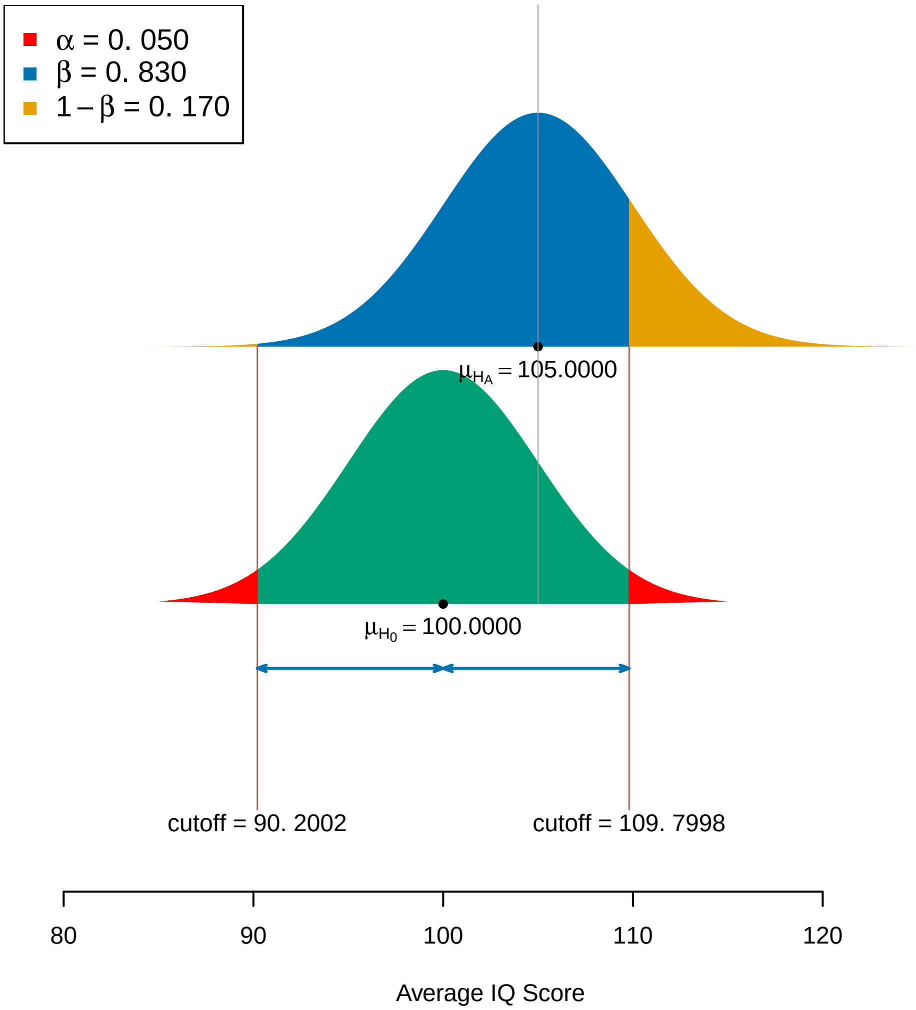
And the power is the sum of these two probabilities:

```
p_left + p_right
```

```
#  [1] 0.170075
```

As show in the following figure:

```
power_plot(n = 9, s = 15, mu0 = 100, mha = 105,
           cutoff = qnorm(c(0.025, 0.975), 100, 15 / sqrt(9)),
           alternative = "equal", xlab = "Average IQ Score")
```



2. Bias in step counters

Following the study by Case et al., JAMA, 2015, suppose we wished to assess, via a formal statistical test, whether (at an *population*, rather than an individual, level) a step-counting device or app is unbiased (H_0) or under-counts (H_A). Suppose we will do so the way Case et al. did, but measuring n persons just once each. We observe the device count when the true count on the treadmill reaches 500.

a. Using a planned sample size of $n = 25$, and $\sigma = 60$ steps as a pre-study best-guess as to the s that might be observed in them, calculate the critical value at $\alpha = 0.01$.

```
qnorm(p = 0.01, mean = 500, sd = 60/sqrt(25))
```

```
#  [1] 472.0838
```

The critical value, i.e., the mean step counts to reject the null is 472.08 steps.

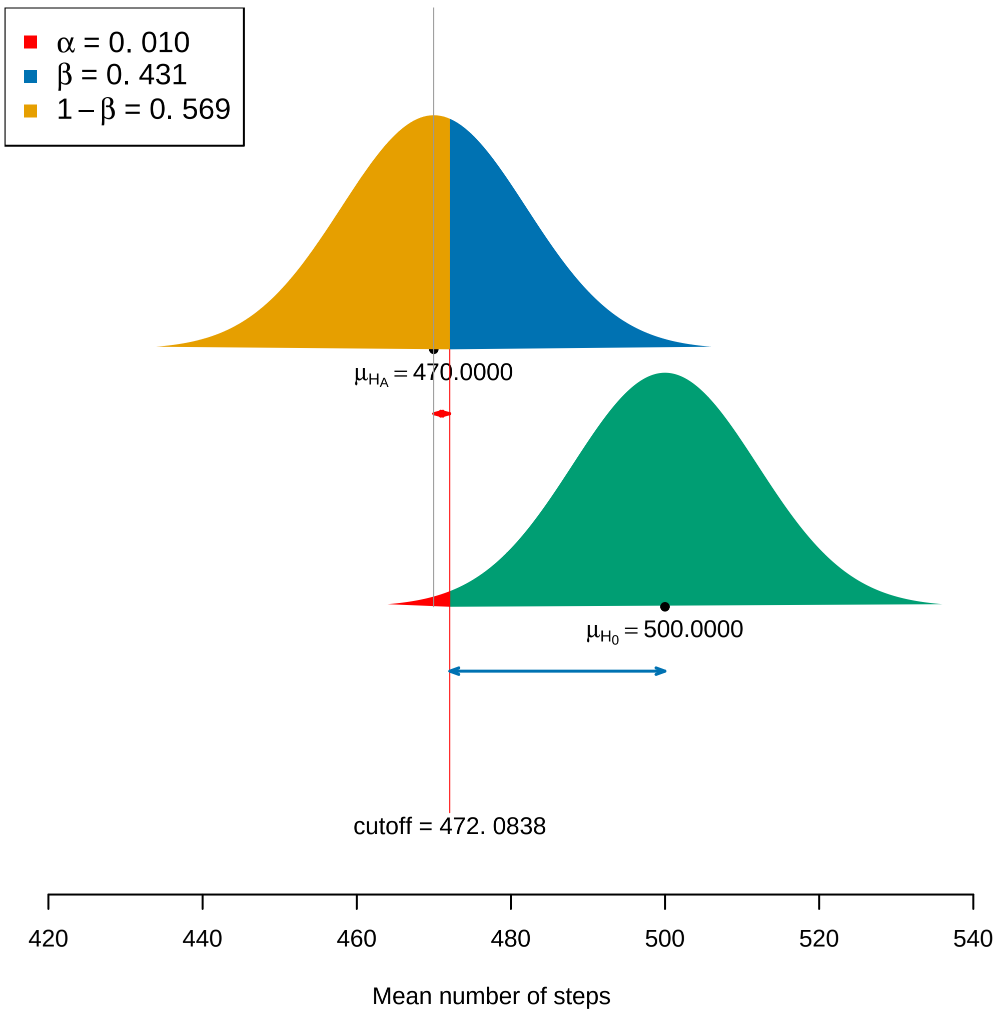
b. Now imagine that the mean would not be the null 500, but $\mu = 470$. Calculate the probability that the mean in the sample of 25 will be less than this critical value. Use the same σ for the alternative that you used for the null. What is this probability called?

```
critical_z <- qnorm(p = 0.01, mean = 500, sd = 60/sqrt(25))
pnorm(q = critical_z, mean = 470, sd = 60/sqrt(25), lower.tail = TRUE)
```

```
#  [1] 0.5689306
```

The probability of getting a sample of size 25 with a mean less than the critical value of 472.08 steps is 0.5689306. This is the statistical power to detect a mean difference of 30 fewer steps.

```
power_plot(n = 25, s = 60, mu0 = 500, mha = 470,
           cutoff = critical_z,
           alternative = "less", xlab = "Mean number of steps")
```



c. Determine the sample size required to detect a 30 step mean decrease in steps with 80% power using a 1% level of significance. Plot the null and alternative distributions in a diagram using the `plot_power` function.

Under the null hypothesis, we know that the mean step count has to be 2.32 standard errors of the mean away from $\mu_0 = 500$ in order to reject the null hypothesis at an $\alpha = 0.01$. 2.32 is the z value such that there is 0.01 area in the left tail of the null distribution:

```r
qnorm(p = 0.01, lower.tail = TRUE)
```

```
#  [1] -2.326348
```

Under the alternative hypothesis, we know that the mean step count has to be 0.84 standard errors of the mean (SEM) away from $\mu_A = 470$ such that there is 20% area in the right tail of the alternative hypothesis distribution:

```r
qnorm(p = 0.20, lower.tail = FALSE)
```

```
#  [1] 0.8416212
```

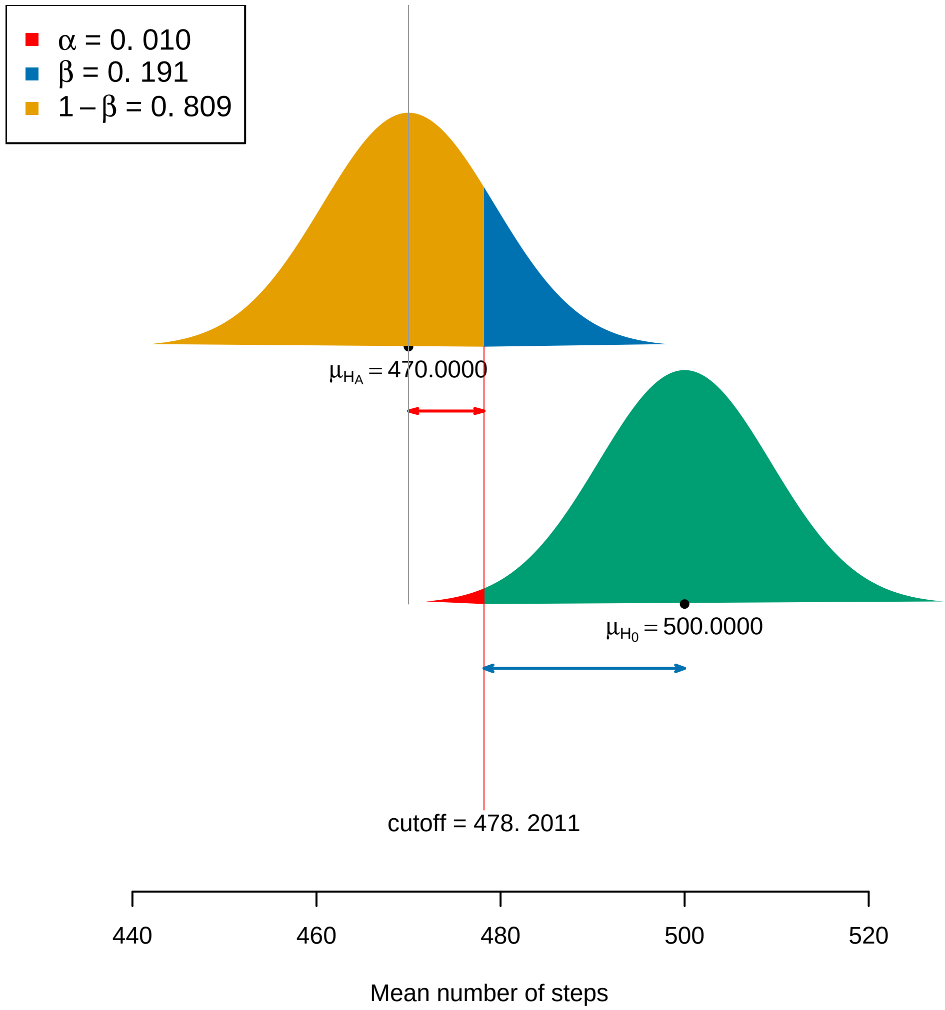
We know that the distance between $\mu_0 = 500$ and $\mu_A = 470$ must be equal to 0.84SEM + 2.32SEM. Note that although the quantile calculated under the null is negative, since we are dealing with distance, we use the absolute value. This is the balancing equation:

$$
\begin{aligned}
\Delta &= 0.84 \times SEM + 2.32 \times SEM \\
\Delta &= (0.84 + 2.32) SEM \\
&= (0.84 + 2.32) \frac{\sigma}{\sqrt{n}} \\
\sqrt{n} &= (0.84 + 2.32) \frac{\sigma}{\Delta} \\
n &= (0.84 + 2.32)^2 \left(\frac{\sigma}{\Delta}\right)^2 \\
&= (0.84 + 2.32)^2 \left(\frac{60}{30}\right)^2 \\
&= 40.14
\end{aligned}
$$

Therefore we need 41 subjects to detect a 30 step mean decrease in steps with 80% power using a 1% level of significance.

```r
source("https://raw.githubusercontent.com/sahirbhatnagar/EPIB607/master/inst/code/plot_null_alt.R

power_plot(n = 41,
           s = 60,
           mu0 = 500,
           mha = 470,
           cutoff = qnorm(0.01, mean = 500, sd= 60/sqrt(41), lower.tail = TRUE),
           alternative = "less",
           xlab = "Mean number of steps ")
```


α = 0. 010
β = 0. 191
1 − β = 0. 809
$\mu_{H_A} = 470.0000$
$\mu_{H_0} = 500.0000$
cutoff = 478. 2011
440
460
480
500
520
Mean number of steps