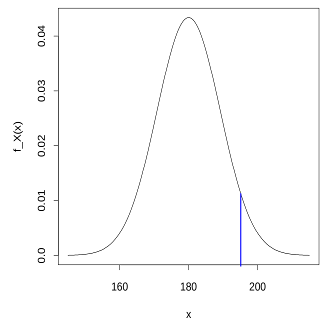
Consider the distribution of serum cholesterol levels for all 16-19 year old males in the US. Suppose we know that the mean of this population is 180 and the standard deviation of this population is 46. You will recall that we have talked much about the distribution of cholesterol levels for U.S. men aged 20-74. Suppose we know, as before, that the population standard deviation is also 46; but suppose we don't know the population mean. We expect that this population mean is larger than the mean in the younger population. Let us conduct a test of this hypothesis.

$$H_0 : \mu \leq 180 \text{ and } H_A : \mu > 180$$

We can afford to collect a sample of 25 males. How big does the sample mean need to be so that we reject the null ? Suppose the probability of making a type I error is set to 0.05 ($\alpha = 0.05$).



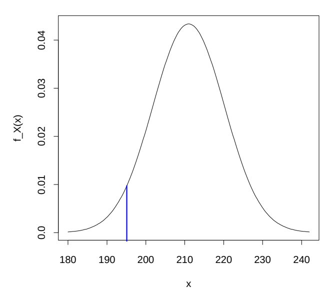
Shown is the distribution of the means of samples of size 25. Samples are taken from a population with mean 180 and standard deviation 46. By the CLT, we know that the distribution of sample means has mean 180 and standard deviation $46/\sqrt{25}$. Sometimes (5 % of the time) we see sample means greater than 195.1.

What is β ? Recall the the type II error probability, β, is the probability of failing to reject the null when it is in fact false. Let's calculate β. To do this, we must specify the way in which the null is false. Suppose the true underlying mean of the older population is 211.

Recall that we figured out that we will reject the null when the sample mean is bigger than or equal to 195.1. Thus, we fail to reject the null when the sample mean is less than 195.1.

So, the question "how often will we fail to reject $H_0 : \mu \leq 180$ given that the true population mean is 211 ?" is equivalent to asking "how often will we see a sample mean less than 195.1 (calculated from a sample of size 25) from a population with true mean 211 ?"

The question "how often will we fail to reject $H_0 : \mu \leq 180$ given that the true population mean is 211 ?" is equivalent to asking "how often will we see a sample mean less than 195.1 (calculated from a sample of size 25) from a population with true mean 211 ?"



Shown is the distribution of the means of samples of size 25. Samples are taken from a population with mean 211 and standard deviation 46. By the CLT, we know that the distribution of sample means has mean 211 and standard deviation $46/\sqrt{25}$. Sometimes (4.2 % of the time) we see sample means less than 195.1.

Recall:

A **type II error** occurs when the null hypothesis is not rejected when it is, in fact, false. The probability of a type II error is denoted by β. Here, $\beta = 0.042$

Note that β depends on the true population mean. It is found by considering the situation in which H_0 is false. H_0 could be false in an infinite number of ways (the true population mean could be anything bigger than 180 in this example).

Power:

If β is the probability of making a type II error, then $1 - \beta$ is the power of the hypothesis. It is the probability of rejecting H_0 given that it is false.

In the last example, the power is $1 - 0.042 = 0.958$

$$\begin{aligned} power &= P\left(\text{reject } \mu \leq 180 | \mu = 211\right) \\ &= P(\bar{X} \geq 195.1 | \mu = 211) \\ &= P(Z \geq -1.73) \\ &= 1 - P(Z < 1.73) \\ &= 1 - 0.042 \\ &= 0.958 \end{aligned}$$

Note that the power $(1 - \beta)$ would have been different if the true mean was different than 211. In practice, it is common to set the type I error rate (α), calculate the power for different values of the true population mean (since you never know these), and plot a power curve.

What is β if the true underlying mean is not 211, but is 200 ?

Sample Size:

We just now figured out how to calculate the power of a test conducted at some level of α for sample size n. Often, at the onset of a study, an investigator would like to know the sample size required to achieve a particular level of power for fixed α. Consider our last example where we are performing a test to make inferences about μ, the true underlying population mean of serum cholesterol levels in 20-74 year old U.S. men ($\sigma = 46$). Suppose we want $\beta = 0.05$ so that power = 95% IF the true underlying mean is 211. Suppose also that we want $\alpha = 0.01$. How big should the sample size be ?

Since $\alpha = 0.01$, we reject the null when it is in fact true only 1% of the time.

With $\alpha = 0.01$, we will reject here if $z \geq 2.32$. In other words, we will reject when $\bar{x} \geq 180 + 2.32\frac{46}{\sqrt{n}}$.

Since we want power=0.95, we know that the probability of rejecting when the null is false ($\mu = 211$) must be 0.95.

Recall that we will reject when $\bar{x} \geq 180 + 2.32\frac{46}{\sqrt{n}}$.

$$
\begin{aligned}
P(\bar{X} \geq 180 + 2.32\frac{46}{\sqrt{n}} \mid \mu = 211) &= 0.95 \\
P(\frac{\bar{X} - 211}{\frac{46}{\sqrt{n}}} \geq \frac{180 - 211}{\frac{46}{\sqrt{n}}} + 2.32) &= 0.95
\end{aligned}
$$