Recap inference (Hypothesis testing)

→ A claim is made.

Example: μ is at most 10

$H_0: \mu \le 10 \rightarrow \mu_0$

$H_1: \mu > 10$

→ Take a sample of size n

→ get sample mean $\bar{x}$

sample standard deviation s ← (unless we somewhat know σ)

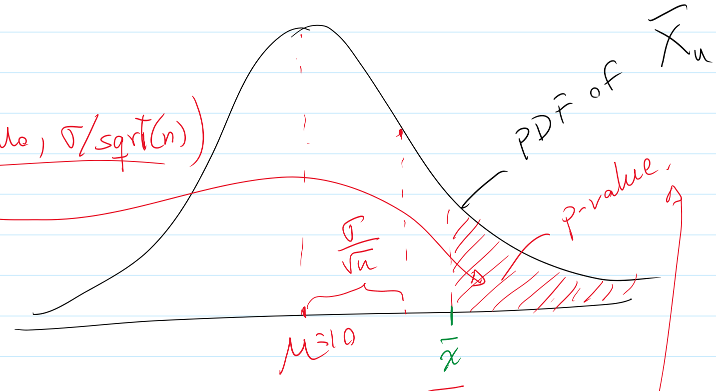
Recall $\bar{X}_n$

$E\bar{X}_n = \mu$

$\mathrm{Var}\,\bar{X}_n = \frac{\sigma^2}{n}$

$SD(\bar{X}_n) = \frac{\sigma}{\sqrt{n}}$

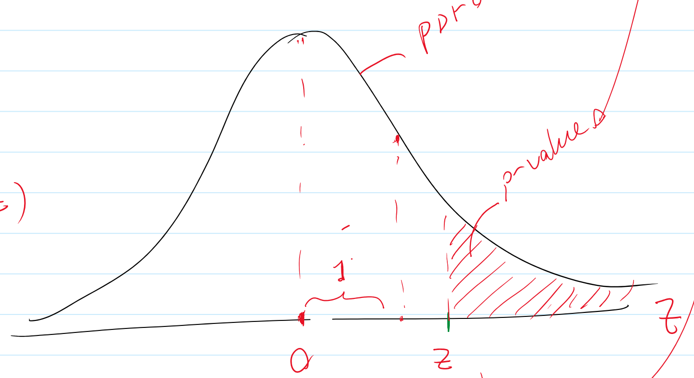
1-pnorm($\bar{x}$, μ_0, σ/sqrt(n))



Let's build our test statistic $z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$

$Z = \frac{\bar{X}_n - \mu_0}{\sigma/\sqrt{n}}$

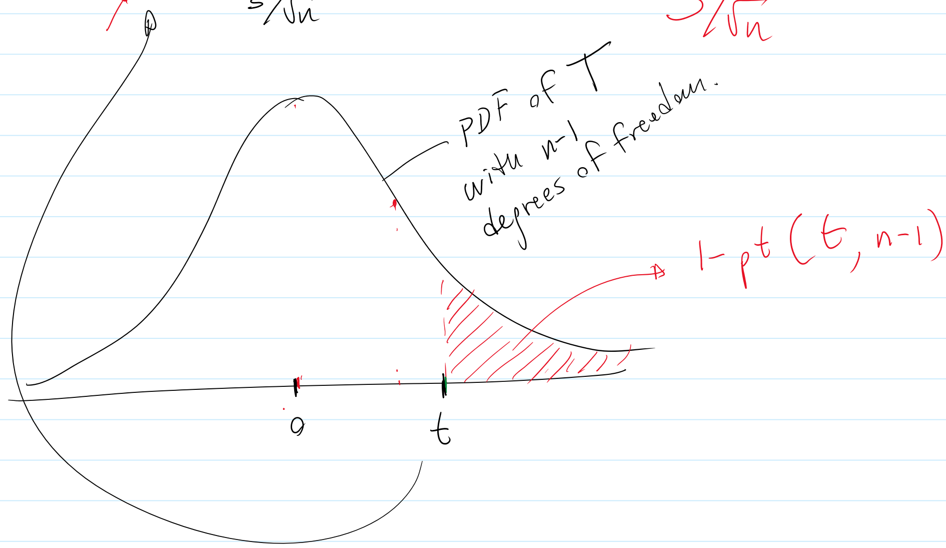
1-pnorm(z)



If we don't know σ, then

$T = \bar{X}_n - \mu_0$

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \qquad T_n = \frac{\bar{X}_n - \mu_0}{S/\sqrt{n}}$$



Examples:

4. The U.S. Food and Drug Administration requires evaporated milk to contain "not less than 23 percent by weight of total milk solids." A company that sells evaporated milk is sued by a group of consumers who are concerned that the company's product does not meet FDA standards. The two parties agree to binding arbitration. If the consumers win, the company will pay damages and enhance its product; if the company wins, then the consumers will issue a public apology.

   To resolve the dispute, the arbiter commissions a neutral study in which the percent by weight of total milk solids will be measured in a random sample of $n = 225$ packages produced by the company. Both parties agree to a standard of proof ($\alpha = 0.05$), but they disagree on which party should bear the burden of proof.

   (a) State appropriate null and alternative hypotheses from the perspective of the consumers.

   (b) State appropriate null and alternative hypotheses from the perspective of the company.

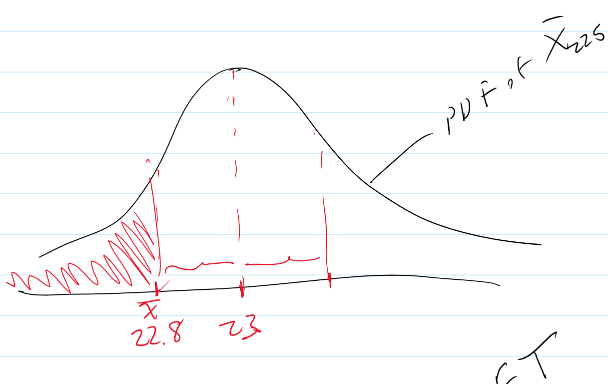
   (c) Suppose that the random sample reveals a sample mean of $\bar{x} = 22.8$ percent with a sample standard deviation of $s = 3$ percent. Compute t, the value of the test statistic.

a) $H_0: \mu \le 23$

$H_1: \mu > 23$

b) $\boxed{H_0: \mu \ge 23 \qquad H_1: \mu < 23}$

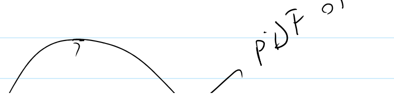
b)



Standard error

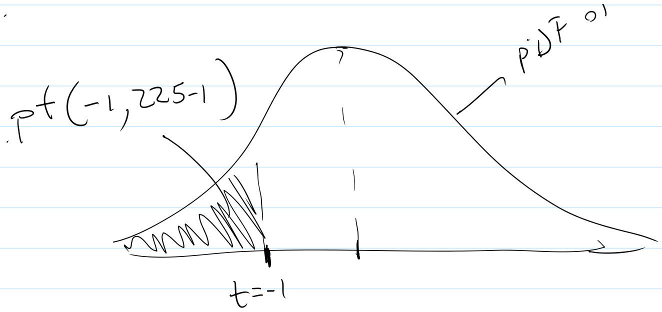
$SD\ \bar{X}_n = \frac{\sigma}{\sqrt{n}}$

std. error $\bar{X}_n = \frac{S}{\sqrt{n}} = \frac{3}{\sqrt{225}} = 0.2$



((, 225-1)

$t = \frac{22.8 - 23}{} = -1$



$$t = \frac{22.8 - 23}{3/\sqrt{225}} = -1$$

```
> mu0 = 23
> n = 225
> xbar = 22.8
> s = 3
> t.v = (xbar - mu0)/(s/sqrt(n))
> t.v
[1] -1
> pt(t.v, n-1)
[1] 0.15919
```

→ p-value.

$\alpha = 0.05$

Do not reject H_0

Fail to reject H_0

~~Accept H_0~~

For Thursday:

5. It is thought that human influenza viruses originate in birds. It is quite possible that, several years ago, a human influenza pandemic was averted by slaughtering 1.5 million chickens brought to market in Hong Kong. Because it is impossible to test each chicken individually, such decisions are based on samples. Suppose that a boy has already died of a bird flu virus apparently contracted from a chicken. Several diseased chickens have already been identified. The health officials would prefer to err on the side of caution and destroy all chickens that might be infected; the farmers do not want this to happen unless it is absolutely necessary. Suppose that both the farmers and the health officals agree that all chickens should be destroyed if more than 2 percent of the population is diseased. A random sample of $n = 1000$ chickens reveals 40 diseased chickens.

   (a) Let $X_i = 1$ if chicken i is diseased and $X_i = 0$ if it is not. Assume that $X_1, \ldots, X_n \sim P$. To what family of probability distributions does P belong? What population parameter indexes this family? Use this parameter to state formulas for $\mu = EX_i$ and $\sigma^2 = \operatorname{Var} X_i$.

   (b) State appropriate null and alternative hypotheses from the perspective of the health officials.