# SYSTEMS 302

# LECTURE 19

- **GENERAL *t*-TESTS**
  - **Summary of *t*-Tests**
  - **P-Values for *t*-Tests**
- **TWO-SAMPLE TESTS**
  - **Difference-between-Means Tests**
  - **Population-Proportions Tests**
  - **Paired *t*-Tests**
- **For next time:**
  - **Devore, Sections 9.2-9.4**

# STOPPING DISTANCE PROBLEM

**To evaluate the safety features of a new surface composition for roads, tests are run on the mean (full-brake) stopping distances for standard passenger vehicles using this surface.**

**DATA:** **A sample of $n = 30$ vehicles are tested for stopping distance at 60 *mph*, yielding data $(x_1,..,x_n)$ with sample mean, $\overline{x}_n = 285.35$ , and standard deviation, $s_n = 13.82$.**

**DECISION RULE:** **Accept new surface only if at least 95% confident that the mean stopping distance, $\mu$, does not exceed 300 *ft* .**

## Hypothesis Formulation:

$$\left.\begin{array}{l} H_0 : \mu \geq 300 \\ H_a : \mu < 300 \end{array}\right\} \text{ accept surface iff } H_0 \text{ is rejected}$$

**Type I Error:** **Accept surface when actually unsafe**

**Significance Level:** $\alpha = .05$

# HYPOTHESIS TESTS WITH VARIANCE UNKNOWN

## TWO-SIDED

$$H_0 : \mu = \mu_0$$
$$H_a : \mu \neq \mu_0$$

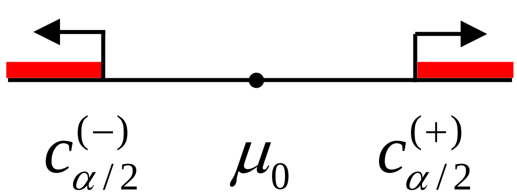
**Choose $\alpha$-error**

**Sampling Error** $= t_{\alpha/2,n-1} \frac{s_n}{\sqrt{n}}$

**Rejection Region for $H_0$:**

$$\left| \bar{X}_n - \mu_0 \right| \geq t_{\alpha/2,n-1} \frac{s_n}{\sqrt{n}}$$

$$c_{\alpha/2} = \mu_0 \pm t_{\alpha/2,n-1} \frac{s_n}{\sqrt{n}}$$




## ONE-SIDED

$$H_0 : \mu = \mu_0$$
$$H_a : \mu < \mu_0$$

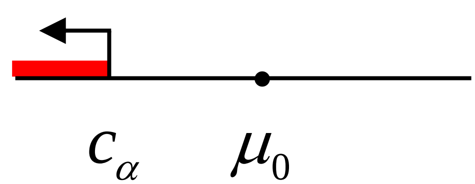
**Choose $\alpha$-error**

**Sampling Error** $= t_{\alpha,n-1} \frac{s_n}{\sqrt{n}}$

**Rejection Region for $H_0$:**

$$\bar{X}_n \leq \mu_0 - t_{\alpha,n-1} \frac{s_n}{\sqrt{n}}$$

$$c_{\alpha} = \mu_0 - t_{\alpha,n-1} \frac{s_n}{\sqrt{n}}$$

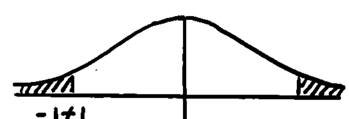
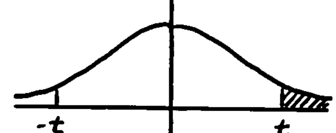
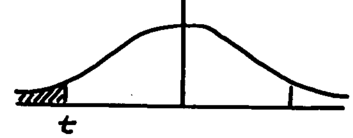


# P-VALUES FOR THE $t$-DISTRIBUTION

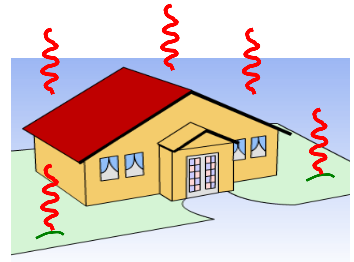
**For any given sample $(x_1,..,x_n)$ to test a null hypothesis $H_0 : \mu = \mu_0$ with unknown variance, we denote the standardized sample mean by**

$$t = \frac{\bar{x}_n - \mu_0}{s_n / \sqrt{n}} \qquad \textbf{(1)}$$

**then the P-value for the test is given by**

| TEST | P-VALUE | |
|---|---|---|
| **Two-Tailed** | $2 \cdot T_{n-1}\left(-\lvert t \rvert\right)$ |   |
| **Upper-Tailed** | $T_{n-1}(-t)$ |   |
| **Lower-Tailed** | $T_{n-1}(t)$ |   |

# RADON PROBLEM



**Radon is the second-leading cause of lung cancer (after smoking). It seeps from the ground into homes, and represents a serious health hazard. A study was done in 1991 to determine the possible relation between indoor radon levels and childhood cancer.**

**DATA: A target sample of radon levels,** $R_{1i}, i = 1,..,n_1$**, was measured in** $n_1 = 42$ **homes with childhood cancer. A control sample,** $R_{2i}, i = 1,..,n_2$**, of** $n_2 = 39$ **similar homes without cancer was also taken. The two sample means** $\bar{R}_1 = 22.81$ **and** $\bar{R}_2 = 19.15$ **suggest a possible 'radon effect'.**

**QUESTION: Is the mean radon level,** $\mu_1$**, in cancer households significantly higher than the mean ,** $\mu_2$**, in cancer-free households?**

**APPROACH: This question can be answered by the following one-sided test:**

$$\left.\begin{array}{l} H_0 : \mu_1 - \mu_2 \le 0 \\ H_a : \mu_1 - \mu_2 > 0 \end{array}\right\}$$

**conclude 'radon effect' if** $H_0$ **rejected**

**Type I Error: Wrongly conclude a 'radon effect'.**

# TEST STATISTIC FOR DIFFERENCES BETWEEN MEANS

**Given independent random samples $(X_{1i} : i = 1,..,n_1)$ and $(X_{2i} : i = 1,..,n_2)$ from two populations with sample means $(\bar{X}_1, \bar{X}_2)$ and sample standard deviations $(s_1, s_2)$, if the test statistic for the true mean difference, $\mu_1 - \mu_2$, between these populations is denoted by**

$$Z_{12} = \frac{\left(\bar{X}_1 - \bar{X}_2\right) - (\mu_1 - \mu_2)}{\sqrt{\left(s_1^2 / n_1\right) + \left(s_2^2 / n_2\right)}}$$

**then for sample sizes $n_1 > 40$ and $n_2 > 40$, one may use the normal approximation, $Z_{12} \sim N(0,1)$. Otherwise, one should use the *t*-approximation, $Z_{12} \sim T_{df}$ , with degrees of freedom given by**

$$df = \frac{(v_1 + v_2)^2}{\dfrac{(v_1)^2}{n_1 - 1} + \dfrac{(v_2)^2}{n_2 - 1}}$$

**where $v_i = s_i^2 / n_i$ , $i = 1,2$ are the *estimated mean variances*.**