Example (AJPH 1999; 89: 1692)

Table 1 in the article by Sargent JD et al provides data on blood lead levels for samples from two communities.

		Providence (1)		Worcester (2)
n		136		153
$\bar{x}$		6.7		5.4
SD		2.06		1.10

Test the hypothesis that $\sigma_1^2 = 6.0$

Test the hypothesis that $\sigma_1^2 = \sigma_2^2$

Test the hypothesis that $\mu_1 = \mu_2$

Module 27: Two Sample t-tests With Unequal Variances

This module shows how to test the hypothesis that two population means are equal when there is evidence that the requirement that the two populations have the same variance is not met.

The General Situation

Earlier we tested hypotheses about two population means, based on data from two independent samples under the assumption that the two populations had the same variance.

For this situation, we calculated the pooled estimate of the common variance with:

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 - 1) + (n_2 - 1)}$$

We then tested the null hypothesis:

$$H_0: \mu_1 = \mu_2 \quad vs \qquad H_1: \mu_1 \neq \mu_2$$

with the test statistic:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

What if $\sigma_1^2 \neq \sigma_2^2$

If the two population variances are not the same, then they are estimated separately and:

$$Var(\bar{x}_1 - \bar{x}_2) = \left(\frac{\sigma_1^{\,2}}{n_1} + \frac{\sigma_2^{\,2}}{n_2} \right)$$

is estimated by:

$$\left(\frac{s_1^{\,2}}{n_1} + \frac{s_2^{\,2}}{n_2} \right)$$

When we have two independent random samples from two different populations and there is evidence that the two population variances differ, ie $\sigma_1^2 \neq \sigma_2^2$, then we can base a test about the difference between the population means μ_1 and μ_2 on the following:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \approx t(f)$$

To deal with $\sigma_1^2 \neq \sigma_2^2$, we are adjusting the degrees of freedom in order to obtain a better approximation than would otherwise be the case. This adjustment requires that we calculate:

$$f = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 + 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 + 1}} - 2$$

Example

Independent random samples were taken from populations of histidine levels for males and females. The measurements were approximately normally distributed.

The statistics from the two samples were:

		Males		Females
n		5		10
$\bar{x}$		300.8		153.2
s^2		15,291.7		2,484.62

Test the hypothesis H_0: $\mu_M = \mu_F$ vs H_1: $\mu_M \neq \mu_F$

1. *The hypothesis:* H_0: $\mu_M = \mu_F$ vs H_1: $\mu_M \neq \mu_F$
2. *The assumptions:* Independent samples, normal distributions
3. *The α-level:* $\alpha = 0.05$
4. *The test statistic:*

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \approx t(f)$$

5. *The critical region:* Reject if t is not between $\pm t_{0.975}(f)$,

where

$$f = \frac{\left[\frac{15{,}291.7}{5} + \frac{2{,}484.6}{10}\right]^2}{\frac{\left[\frac{15{,}291.7}{5}\right]^2}{5+1} + \frac{\left[\frac{2{,}484.6}{10}\right]^2}{10+1}} - 2$$

so that $f = 6.99 - 2 = 4.99$. We will use $f = 5$,

and $t_{0.975}(5) = 2.57058$.

6. *The result:*

$$t = \frac{300.8 - 153.2}{\sqrt{\frac{15,291.7}{5} + \frac{2,486.6}{10}}}$$

$$t = \frac{147.6}{57.50} = 2.57$$

7. *The conclusion:* The value calculated for t above is very close to the cut point so that $P \approx 0.05$

99% Confidence Interval

$$C\left[\bar{x}_1 - \bar{x}_2 - t_{1-\alpha/2}(f)\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} < \mu_1 - \mu_2 < \bar{x}_1 - \bar{x}_2 + t_{1-\alpha/2}(f)\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}\right] = 0.99$$

for $n_1 = 5$, $\bar{x}_1 = 300.8$, $s_1^2 = 15{,}291.70$

$n_2 = 10$, $\bar{x}_2 = 153.2$, $s_2^2 = 2484.62$

for which $\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2} = 3{,}306.8$

$f = 5$, $t_{0.995}(5) = 4.0321$

We have

$$C\left[3008-1532-4.0321\sqrt{3,3068} < \mu_1-\mu_2 < 3008-1532+4.0321\sqrt{3,3068}\right]=0.99$$

$$C\left[147.6-231.87 < \mu_1-\mu_2 < 147.6+231.87\right]=0.99$$

$$C\left[-84.27 < \mu_1-\mu_2 < 379.47\right]=0.99$$

The Association Between State Housing Policy and Lead Poisoning in Children

James D. Sargent, MD, Madeline Dalton, PhD, Eugene Demidenko, PhD, Peter Simon, MD, and Robert Z. Klein, MD

ABSTRACT

Objectives. This study examined the effect of an active program of household lead paint hazard abatement, applied over 22 years, on childhood lead poisoning in Massachusetts.

Methods. A small areas analysis was used to compare screening blood lead levels of children in Worcester County, Mass (n = 27 590), with those in Providence County, RI (n = 19 071). Data were collapsed according to census tract.

Results. The percentage of children with lead poisoning (blood lead level ≥ 20 μg/dL [Pc20]) was, on average, 3 times higher in Providence County census tracts (3.2% vs 0.9% in Worcester County census tracts, $P < .0001$), despite similar percentages of pre-1950s housing in both counties. The ratio of Pc20 in Providence vs Worcester County census tracts was 2.2 (95% confidence interval = 1.8, 2.7), after adjustment for differences in housing, sociodemographic, and screening characteristics. This estimate was robust to alternative regression methods and sensitivity analyses.

Conclusions. Massachusetts policy, which requires lead paint abatement of children's homes and places liability for lead paint poisoning on property owners, may have substantially reduced childhood lead poisoning in that state. (*Am J Public Health.* 1999;89:1690–1695)

Source: Am J Public Health, 1996;Oct:1436

TABLE 1—Mean Values for Aggregate Measures of Lead Exposure in Children Aged 0 to 5 Years Screened in Providence County and Worcester County Census Tracts (CTs)

	Mean ± SD	
Aggregate Measure	Providence (n = 136 CTs)	Worcester (n = 153 CTs)
Blood lead level ≥ 10 μg/dL, %	20.2 ± 12.7	10.3 ± 7.0
Blood lead level ≥ 20 μg/dL, %	3.2 ± 3.6	0.9 ± 1.10
Blood lead level ≥ 30 μg/dL, %	0.8 ± 1.1	0.2 ± 0.51
Arithmetic mean blood level, μg/dL	6.7 ± 2.06	5.4 ± 1.10
Geo mean blood level, μg/dL	6.4	5.3