Fuzzy Statistics

Dr. Mohammed Jasim Mohammed
Estimate $\mu_1 - \mu_2$

Chapter 7
7.1 Estimate $\mu_1 - \mu_2$, Variance Known

We have two populations: Pop I and Pop II. Pop I is normally distributed with unknown mean $\mu_1$ and known variance $\sigma_1^2$. Pop II is also normally distributed with unknown mean $\mu_2$ but known variance $\sigma_1^2$. We wish to construct a fuzzy estimator for $\mu_1 - \mu_2$. We collect a random sample of size $n_1$ from Pop I and let $\bar{x}_1$ be the mean for this data. We also gather a random sample of size $n_2$ from Pop II and $\bar{x}_2$ is the mean for the second sample. We assume these two random samples are independent.
7.1 Estimate $\mu_1 - \mu_2$ , Variance Known

Now $\bar{x}_1 - \bar{x}_2$ is normally distributed with mean $\mu_1 - \mu_2$ and standard deviation $\sigma_0 = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$. Then as in Section 3.3 of Chapter 3, $(1 - \beta)\% 100$ confidence interval for $\mu_1 - \mu_2$, is

\[
\bar{x}_1 - \bar{x}_2 - z_{\beta/2} \sigma_0, \bar{x}_1 - \bar{x}_2 + z_{\beta/2} \sigma_0
\]

we place these confidence intervals one on top of another to build our fuzzy estimator $\tilde{\mu}_{12}$ for $\mu_1 - \mu_2$. 
Example 7.1.1

Assume that: (1) $n_1 = 15, \bar{x}_1 = 70.1, \sigma_1^2 = 6$; and (2) $n_2 = 8, \bar{x}_2 = 75.3, \sigma_2^2 = 4$. Then equation (7.1) becomes

$$[-5.2 - 0.9487z_{\beta/2}, -5.2 + 0.9487z_{\beta/2}]$$
Figure 7.1: Fuzzy Estimator $\mu_{12}$ in Example 7.1.1, $0.01 \leq \beta \leq 1$
8.1 Estimate $\mu_1 - \mu_2$, Variances Unknown: Large Samples

We assume that $n_1 > 30, n_2 > 30$. Let $S_1^2, S_2^2$ be the sample variance calculated from the data acquired from Pop I (Pop II). With large samples we may use the normal approximation and a $(1 - \beta)\%100$ confidence interval for $\mu_1 - \mu_2$ is (Section 7.3 in [1]).

$$\bar{x}_1 - \bar{x}_2 - z_{\beta/2}s_0, \bar{x}_1 - \bar{x}_2 + z_{\beta/2}s_0.$$ 

Where $s_0 = \sqrt{s_1^2/n_1 + s_2^2/n_2}$. Put these confidence intervals together to obtain a fuzzy estimator $\bar{\mu}_{12}$ for the difference of the means. The results are similar to those in Chapter 7.
8.2 Estimate $\mu_1 - \mu_2$, Variances Unknown: Small Samples

Here we have two cases: (1) if we may assume that the variances are equal; or (2) the variances are not equal. We assume that $n_1 < 30, n_2 < 30$. Let $S_1^2, S_2^2$. 
8.3.1 Estimate $\mu_1 - \mu_2$, Variances Unknown: Small Samples: Equal Variances

We have $\sigma_1^2 = \sigma_2^2 = \sigma^2$. Define $S_p$, the pooled estimator of the common variance, as,

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

Let $s_* = s_p \sqrt{1/n_1 + 1/n_2}$. Then it is known that

$$T = \frac{(x_1 - x_2) - (\mu_1 - \mu_2)}{s_*}$$
8.3.1 Estimate $\mu_1 - \mu_2$, Variances Unknown: Small Samples: Equal Variances

has a (Student's) $t$-distribution with $n_1 + n_2 - 2$ degrees of freedom. Then

$$P(-t_{\beta/2} \leq T \leq t_{\beta/2}) = 1 - \beta.$$ 

Solve the inequality for $\mu_1 - \mu_2$ and we find the $(1 - \beta)\%100$ confidence interval for $\mu_1 - \mu_2$.

$$[\bar{x}_1 - \bar{x}_2 - t_{\beta/2}s^*, \bar{x}_1 - \bar{x}_2 + t_{\beta/2}s^*].$$

We place these confidence intervals one on top of another, as in Chapter 3, to get our fuzzy estimator $\bar{\mu}_{12}$. 
Example 8.3.1.1

Assume that the variances in the two populations are equal. Let the data be: (1) $n_1 = 15, \bar{x}_1 = 70.1, s_1 = 6$ and $n_2 = 8, \bar{x}_2 = 75.3, s_2 = $ We compute $S_p = 2.3094$. Then

$$[-5.2 - 1.0110t_{\beta/2}, -5.2 + 1.0110t_{\beta/2}].$$

The degrees of freedom is 21. if $0.01 \leq \beta \leq 1$, then the graph of the fuzzy estimator $\tilde{\mu}_{12}$, is in Figure 8.1.
Example 8.3.1.1

Figure 8.1: Fuzzy Estimator $\bar{\mu}_{12}$ in Example 8.3.1.1, $0.01 \leq \beta \leq 1$
8.3.2 Estimate $\mu_1$-$\mu_2$ , Variances Unknown: Small Samples: Unequal Variances

We have $\sigma_1^2 \neq \sigma_2^2$. Define $S_p$ , It is known that

$$T = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_0}.$$ 

is approximately (Student's) $t$ distributed with $r$ degrees of freedom. We find the degrees of freedom ($r$) by rounding up to the nearest integer the following expression.
8.3.2 Estimate μ₁ - μ₂, Variances Unknown: Small Samples: Unequal Variances

\[
\frac{(A + B)^2}{n_1 - 1} + \frac{B^2}{n_2 - 1}
\]

Where \( A = s_1^2/n_1 \) and \( B = s_2^2/n_2 \). Let

\[
T = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_\bar{x}}
\]

Then

\[
P(-t_{\beta/2} \leq T \leq t_{\beta/2}) \approx 1 - \beta.
\]

Solve the inequality for \( \mu_1 - \mu_2 \) and we find the (1-\( \beta \))%100 confidence interval for \( \mu_1 - \mu_2 \).
8.3.2 Estimate $\mu_1 - \mu_2$, Variances Unknown: Small Samples: Unequal Variances

$$[\bar{x}_1 - \bar{x}_2 - t_{\beta/2} s_0, \bar{x}_1 - \bar{x}_2 + t_{\beta/2} s_0].$$

The t distribution has r degrees of freedom. We place these confidence intervals one on top of another, as in Chapter 3, to get our fuzzy estimator $\bar{\mu}_{12}$. 

Dr. Mohammed Jasim
Example 8.3.2.1

Let us use the same data as in Example 8.3.1.1 except now we do not assume the variances are the same. The data are:

$(1) \quad n_1 = 15, \bar{x}_1 = 70.1, s_1 = 6 \text{ and } n_2 = 8, \bar{x}_2 = 75.3, s_2 = 8$.

We compute $S_p = 2.3094$.

An approximate $(1 - \beta)\%100$ confidence interval for the difference of the means is

$$[-5.2 - (0.9487)t_{\beta/2}, -5.2 + (0.9487)t_{\beta/2}],$$
Example 8.3.2.1

The term $S_0$ was defined in Section 8.2 and we computed it as 0.90487. Also, we determined that the degrees of freedom is $r = 18$. If $0.01 \leq \beta \leq 1$, then the graph of the fuzzy estimator $\mu_\_12$, is in Figure 8.2.
Example 8.3.2.1

Figure 8.2: Fuzzy Estimator $\overline{\mu}_{12}$ in Example 8.3.2.1, $0.01 \leq \beta \leq 1$
9.1 Estimate $d=\mu_1-\mu_2$, Matched Pairs

Let $x_1, x_2, ..., x_n$ be the values of a random sample from a population Pop 1. Let $object_i$ or $person_i$, belong to Pop 1 which produced measurement $x_i$ in the random sample, $1 < i < n$. Then, at possibly some later time, we take a second measurement on $object_i$ ($person_i$) and get value $y_i, 1 < i < n$. Then $(x_1, y_1), \cdots (x_n, y_n)$, are n pairs of dependent measurements.
9.1 Estimate $d = \mu_1 - \mu_2$, Matched Pairs

For example, when testing the effectiveness of some treatment for high blood pressure, the $x_i$ are the blood pressure measurements before treatment and the $y_i$ are these measurements on the same person after treatment. The two samples are not independent so we cannot use the results of Chapters 7 or 8.
9.1 Estimate d=µ1- µ2 , Matched Pairs

Let \( d_i = x_i - y_i \), \( 1 < i < n \). Next compute the mean \( \bar{d} \) (crisp number, not fuzzy) and the variance \( S_d^2 \) of the \( d_i \) data. Assume that \( n > 30 \) so we may use the normal approximation; or assume that the \( d_i \) are approximately normally distributed with unknown mean \( \mu_d \) and unknown variance \( \sigma_d^2 \). Then

\[
T = \frac{d - \mu_d}{S_d / \sqrt{n}},
\]
9.1 Estimate \( d = \mu_1 - \mu_2 \), Matched Pairs

has a t distribution with \( n - 1 \) degrees of freedom. It follows that: 

\[
P(-t_{\beta/2} \leq T \leq t_{\beta/2}) = 1 - \beta.
\]

From this it immediately follows that a \( (1-\beta)\%100 \) confidence interval for \( \mu_d \) is:

\[
[\bar{d} - t_{\beta/2} \frac{s_d}{\sqrt{n}}, \bar{d} + t_{\beta/2} \frac{s_d}{\sqrt{n}}].
\]

Now place these confidence intervals given in above equation one on top of another to produce our fuzzy estimator \( \bar{\mu}_d \) of \( \mu_d \).
Example 9.1.1
Consider the paired data in the below Table. This table contains a weeks forecast high temperatures and the actual recorded high values.

<table>
<thead>
<tr>
<th>Forecast (x)</th>
<th>Actual High (y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>68</td>
<td>72</td>
</tr>
<tr>
<td>76</td>
<td>74</td>
</tr>
<tr>
<td>66</td>
<td>62</td>
</tr>
<tr>
<td>72</td>
<td>76</td>
</tr>
<tr>
<td>76</td>
<td>75</td>
</tr>
<tr>
<td>80</td>
<td>78</td>
</tr>
<tr>
<td>71</td>
<td>75</td>
</tr>
</tbody>
</table>
Example 9.1.1

We compute \( d_i = x_i - y_i, 1 < i < 7 \). and then \( \bar{d} = -0.4286, S_d = 3.4572 \).

The \((1-\beta)\%100\) confidence interval are

\[
[-0.4286 - (1.3067)t_{\beta/2}, -0.4286 + (1.3067)t_{\beta/2}].
\]

We graphed the confidence intervals for \(0.01 \leq \beta \leq 1\), and the result is our fuzzy estimator \( \bar{\mu}_d \) of \( \mu_d \) in Figure 9.1.
Example 9.1.1

Figure 9.1: Fuzzy Estimator $\bar{\mu}_d$ in Example 9.1.1, $0.01 \leq \beta \leq 1$