Chapter 20: Comparing Groups

1. **Confidence Interval for the Difference Between Two Proportions:**

**Seat-Belt Example:**

1. Open a new excel sheet.
2. Using the textbook example: sample size for men (n) = 4208, sample size for women (m) = 2763, successes in men = 2777, successes in women = 1363 and construct a 95% confidence interval.

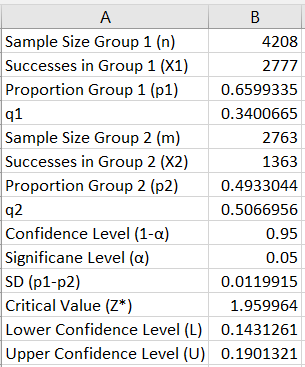
|  |  |
| --- | --- |
| Sample Size Group 1 (n) | = 4208 (given) |
| Successes in Group 1 (X1) | = 2777 (given) |
| Proportion Group 1 (p1) |  |
|  |  |
| Sample Size Group 2 (m) | = 2763 (given) |
| Successes in Group 2 (X2) | = 1363 (given) |
| Proportion Group 2 (p2) |  |
|  |  |
| Confidence Level (1 - α) | = 0.95 (given) |
| Significant level (α) |  |
| SD (p1-p2) |  |
| Critical Values (Z\*) |  |
| Lower Confidence Level (L) |  |
| Upper Confidence Level (U) |  |

1. Use **NORM.S.INV** function to calculate critical values of normal distribution.



where **Probability** is the probability corresponding to the normal distribution.

1. The result is



The 95% confidence interval is (0.143, 0.190)

1. **Testing for the Difference Between Two Proportions:**

**Sleep Example:**

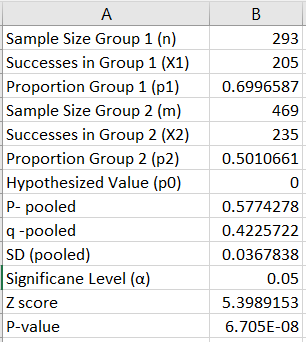
1. Open a new excel sheet.
2. Using the textbook example: sample size for men (n) = 4208, sample size for women (m) = 2763, successes in men = 2777, and successes in women = 1363.

Are the pre-sleep surfing rate or the two groups really different?

|  |  |
| --- | --- |
| Sample Size Group 1 (n) | = 4208 (given) |
| Successes in Group 1 (X1) | = 2777 (given) |
| Proportion Group 1 (p1) |  |
| Sample Size Group 2 (m) | = 2763 (given) |
| Successes in Group 2 (X2) | = 1363 (given) |
| Proportion Group 2 (p2) |  |
| Hypothesized Value (p0) | 0 |
| Pooled sample proportion (p-pooled) |  |
| q (pooled) |  |
| SD (pooled) |  |
| Significant level (α) | = 0.05 (given) |
| Z score |  |
| P-value |  |

Note: the functions were used in chapter 5.

1. The result is



**Note:**

1. If , then use P-value = NORM.S.DIST (z, True)
2. If , then use P-value = 1 - NORM.S.DIST (z, True)
3. If , then use P-value = 2 (1 - NORM.S.DIST (ABS(z), True))

**Note:** **ABS** function returns the absolute value of a number.

1. **Confidence Interval for the Difference Between Two Means:**

**Soup Bowls Example:**

1. Open a new excel sheet.
2. Using the textbook example:

Refilling bowl: sample size (n) = 27, mean = 14.7, and standard deviation = 8.4. Ordinary bowl: sample size (n) = 27, mean = 8.5, and standard deviation = 6.1.

Construct a 95% confidence interval.

|  |  |
| --- | --- |
| Sample Size Group 1 (n) | = 27 (given) |
| Mean Group 1 – Refilling bowl | = 14.7 (given) |
| Standard Deviation Group 1 (SD1) | = 8.4 (given) |
| Sample Size Group 2 (m) | = 27 (given) |
| Mean Group 2 – Ordinary bowl | = 8.5 (given) |
| Standard Deviation Group 2 (SD2) | = 6.1 (given) |
| Standard Deviation of the Difference |  |
| Confidence Level (1 - α) | = 0.95 (given) |
| Significant Level (α) |  |
| Degrees of Freedom (df) |  |
|  | |
| Critical Value (t) |  |
| Lower Confidence Level (L) |  |
| Upper Confidence Level (U) |  |

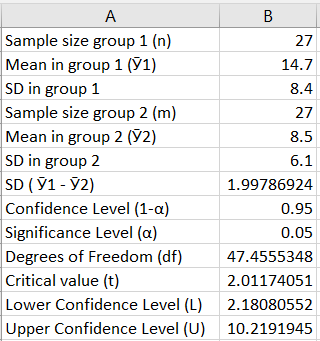
1. Use **T.INV.2T** function to calculate critical values of normal distribution.



where **Probability** is the probability associated with t-distribution.

**Deg\_freedom** The number of degrees of freedom with which to characterize the distribution.

1. The result is



The 95% confidence interval is (2.18, 10.22)

1. **Testing for the Difference Between Two Means:**

**Friend of Stranger Example:**

1. Open a new excel sheet.
2. Using the textbook example:

Buy from friend: sample size (n) = 8, mean = 281.88, and standard deviation = 18.31. Buy from stranger: sample size (n) = 7, mean = 211.43, and standard deviation = 46.43.

Is there a difference in the average price people would offer a friend rather than a stranger?

|  |  |
| --- | --- |
| Sample Size Group 1 (n) | = 8 (given) |
| Mean Group 1 – Refilling bowl | = 281.88 (given) |
| Standard Deviation Group 1 (SD1) | = 18.31 (given) |
| Sample Size Group 2 (m) | = 7 (given) |
| Mean Group 2 – Ordinary bowl | = 211.43 (given) |
| Standard Deviation Group 2 (SD2) | = 46.43(given) |
| Standard Deviation of the Difference |  |
| Hypothesized Value () | = 0 (given) |
| Significant Level (α) | = 0.05 (given) |
| t Score |  |
| Degrees of Freedom (df) |  |
|  | |
| P-Value |  |

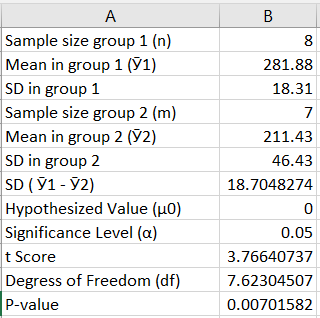
1. Use **T.DIST.2T** function to calculate the p-value.



where **x** is the test statistic (t).

**Deg\_freedom** The number of degrees of freedom with which to characterize the distribution.

1. The result is



**Note:**

1. If , then use P-value = T.DIST (t, df, True)
2. If , then use P-value = T.DIST.RT (t, df)
3. If , then use P-value = T.DIST.2T (t, df)