Confidence Interval Estimation of $\mu_1 - \mu_2$ when $\sigma_1^2 \neq \sigma_2^2$ Not common (Welch–Satterthwaite)

Result: Let $X_{1,1}, X_{1,2}, \cdots, X_{1,n1}$ represent a random sample taken from a population of values with a mean μ_1 and a variance σ_1^2. Additionally, let $X_{2,1}, X_{2,2}, \cdots, X_{2,n2}$ represent a random sample taken from a different population having a mean μ_2 and a variance σ_1^2, where $X_{2,i}$ is the same variable as $X_{1,i}$. A $100(1 - \alpha)\%$ confidence interval estimate for the difference between the two population means $\mu_1 - \mu_2$ is

$$(\bar{X}_1 - \bar{X}_2) \pm T_{\alpha/2,(df)} \sqrt{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)}$$

Where

$$df \approx \left\lfloor \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{S_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{S_2^2}{n_2}\right)^2} \right\rfloor = floor\left(\frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{S_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{S_2^2}{n_2}\right)^2}\right)$$

do not worry About showing This is The df just believe me!

We usually take the floor because it is a conservative estimate, and most people will use discrete t-tables. If you would like more information you can research Behrens–Fisher problem.

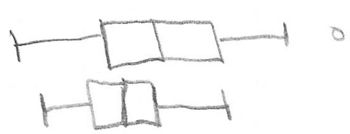
Example 4: In 2008, researchers carried out a study[2] where a random sample of men and a random sample of women, all of whom worked as purchasing managers in Canada and subscribed to the Purchasing Magazine. Each male and female were asked to provide the researchers with their annual salary, in \$1000s. The results of this study were published in the paper "Sex and Salary: A Survey of Purchasing and Supply Professionals". The raw data is provided in the R-code, summarize the data and analyze the implications with 99% confidence.

Use R-code to see ...

	min	Q1	Q2	Q3	n	$\bar{X}$	sd	max
female	49.72	59.19	64.94	67.60	34	63.90	6.29	78.46
male	48.41	63.91	71.22	77.90	58	72.14	10.36	100.17097

male

female.



- 99% CI

$0.2 < \frac{\sigma_F^2}{\sigma_M^2} < 0.7 \Rightarrow \sigma_F^2 < \sigma_M^2$ σ_F^2 more Narrow More Predictable.

but $\sigma_F^2 \neq \sigma_M^2$ Sig. different

Don't pool!

Based on this sample we are 99% Confident the Average difference between female salary and male salary is Between (−12.8, −3.6) Does not contain Zero ⇒ on Average females make significantly less than males.

df = 89.853 ⇒ 89.

$(63.9 - 72.14) \pm 2.6317 \left[\sqrt{\frac{6.29^2}{34} + \frac{10.36^2}{58}}\right]$

[2] http://www.sciencedirect.com/science/article/pii/S1478409208000113