ON THE USE OF THE NON-CENTRAL t-DISTRIBUTION FOR COMPARING PERCENTAGE POINTS OF NORMAL POPULATIONS

By John E. Walsh

Princeton University

1. Introduction. Consider two normal populations with the same variance and means μ and ν respectively. It is well known that confidence intervals and significance tests can be obtained for the difference $\mu - \nu$. Since μ is the 50% point of the first population and ν is the 50% point of the second population, this represents a particular solution of the general problem of obtaining confidence intervals and significance tests for the difference $\theta_\alpha - \varphi_\beta$, where θ_α is the α percent point of the first population and φ_β is the β percent point of the second population. The purpose of this note is to point out that the results of Johnson and Welch [1] for the non-central t-distribution can be used to furnish a solution of the general problem.

2. Analysis. Let A_γ be the γ percent point of the normal population with zero mean and unit variance (i.e. exactly γ% of the population has values less than A_γ). Then if σ is the common standard deviation,

$$\theta_\alpha = \mu + A_\alpha\sigma, \qquad \varphi_\beta = \nu + A_\beta\sigma.$$

Thus

$$\theta_\alpha - \varphi_\beta = (\mu - \nu) + (A_\alpha - A_\beta)\sigma.$$

The non-central t-distribution investigated by Johnson and Welch in [1] is based on the quantity

$$t = (z + \delta)/\sqrt{\chi^2/f},$$

where z has a normal distribution with zero mean and unit variance, δ is a constant, and χ^2 has a χ^2-distribution with f degrees of freedom and is distributed independently of z. Methods and tables are given in [1] whereby a constant $t(f, \delta, \epsilon)$ can be computed having the property that

$$Pr[t > t(f, \delta, \epsilon)] = \epsilon.$$

These relations will be used to obtain confidence intervals for $\theta_\alpha - \varphi_\beta$. The resulting confidence intervals can be used to obtain significance tests for $\theta_\alpha - \varphi_\beta$.

Let $x_1, \cdots, x_n$ be a random sample of size n from the first population while $y_1, \cdots, y_m$ is a random sample of size m from the second population. Then consider

$$\frac{\bar{x} - \bar{y} - (\theta_\alpha - \varphi_\beta)}{\sqrt{\sum_1^n (x_i - \bar{x})^2 + \sum_1^m (y_j - \bar{y})^2}} \cdot \sqrt{\frac{m + n - 2}{\frac{1}{n} + \frac{1}{m}}} = \frac{\left[\frac{\bar{x} - \bar{y} - (\mu - \nu)}{\sigma\sqrt{\frac{1}{n} + \frac{1}{m}}}\right] + \frac{(A_\beta - A_\alpha)}{\sqrt{\frac{1}{n} + \frac{1}{m}}}}{\sqrt{\frac{\sum_1^n (x_i - \bar{x})^2 + \sum_1^m (y_j - \bar{y})^2}{\sigma^2(m + n - 2)}}}.$$

This quantity has a non-central t-distribution with

$$\delta = (A_\beta - A_\alpha) \Big/ \sqrt{\frac{1}{n} + \frac{1}{m}}, \qquad f = m +$$

For notational simplicity let

$$t\left(m + n - 2, \frac{A_\beta - A_\alpha}{\sqrt{\frac{1}{n} + \frac{1}{m}}}, \epsilon\right) = t(\epsilon), \quad \sum_1^n (x_i - \bar{x})^2 = S_1^2, \quad \sum_1^m (y_i - \bar{y})^2 = S_2^2.$$

Then one-sided confidence intervals for $\theta_\alpha - \varphi_\beta$ with confidence coefficient ϵ are given by

$$\theta_\alpha - \varphi_\beta < \bar{x} - \bar{y} - \frac{t(\epsilon)\sqrt{S_1^2 + S_2^2}}{\sqrt{(m + n - 2) \Big/ \left(\frac{1}{n} + \frac{1}{m}\right)}},$$

$$\theta_\alpha - \varphi_\beta > \bar{x} - \bar{y} - \frac{t(1 - \epsilon)\sqrt{S_1^2 + S_2^2}}{\sqrt{(m + n - 2) \Big/ \left(\frac{1}{n} + \frac{1}{m}\right)}}.$$

Two-sided confidence intervals for $\theta_\alpha - \varphi_\beta$ with confidence coefficient

$$1 - (\epsilon_1 + \epsilon_2)$$

are given by

$$\bar{x} - \bar{y} - \frac{t(\epsilon_2)\sqrt{S_1^2 + S_2^2}}{\sqrt{(m + n - 2) \Big/ \left(\frac{1}{n} + \frac{1}{m}\right)}} < \theta_\alpha - \varphi_\beta < \bar{x} - \bar{y} - \frac{t(1 - \epsilon_1)\sqrt{S_1^2 + S_2^2}}{\sqrt{(m + n - 2) \Big/ \left(\frac{1}{n} + \frac{1}{m}\right)}},$$

where $\epsilon_1 + \epsilon_2 < 1$.

REFERENCE

[1] N. L. Johnson and B. L. Welch, "Applications of the non-central t-distribution", *Biometrika*, Vol. 31 (1940), pp. 362–389.