Probability Distributions Useful in Constructing Confidence Intervals and Tests

We can utilize the information we have developed about probability to form confidence intervals, for example. In particular, we shall utilize the following facts:

(a) If $X_1, \ldots X_n$ are iid $N(\mu, \sigma^2)$ then $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$ and this random variable is independent of $\bar{X}$.

(b) If $W_i, i = 1, \ldots, k$ are independent, $W_i \sim \chi^2_{n_i}$ then $\sum_{i=1}^k W_i \sim \chi^2_m$, where $m = \sum_{i=1}^k n_i$.

(c) The t-distribution with m degrees of freedom is the distribution of the random variable $\frac{Z}{\sqrt{W/m}}$, where $Z \sim N(0,1)$ and $W \sim \chi^2_m$ are independent.

(d) If $X \sim N(\mu, \sigma^2)$ then $\frac{X-\mu}{\sigma} \sim N(0,1)$.

(e) The F-distribution with m df for the numerator and n df for the denominator is the probability distribution of the random variable $\frac{U/m}{V/n}$ where $U \sim \chi^2_m$ and $V \sim \chi^2_n$ and U and V are independent.

(f) If $X_1, \ldots, X_n$ are iid from a distribution with mean μ and variance σ^2 then (by the central limit theorem) approximately $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$. If the X's are from a normal population, then the last is exact.

(g) If $X_1, \ldots, X_n$ are iid $\sim N(\mu_X, \sigma^2_X)$ and $Y_1, \ldots, Y_m$ are iid $\sim N(\mu_Y, \sigma^2_Y)$ and the full set $X_1, \ldots, Y_m$ are independent then $\bar{X} - \bar{Y} \sim N(\mu_X - \mu_Y, \frac{\sigma^2_X}{n} + \frac{\sigma^2_Y}{m})$ and S^2_X and S^2_Y are independent of each other and $\bar{X} - \bar{Y}$.

We have, as a consequence, the following

1. If $X_1, \ldots, X_n$ are iid $N(\mu, \sigma^2)$ then by (a) we can use this to "trap" σ^2 for a confidence interval.

2. If $X_1, \ldots, X_n$ are iid $N(\mu, \sigma^2)$ then from (a), (c), and (f), $\frac{\bar{X}-\mu}{s/\sqrt{n}} \sim t_{n-1}$. We can used this to get a confidence interval for μ .

3. From (g), (c), (b), and (a) we know that if $\sigma_X = \sigma_Y$ then

$$\frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{s_p\sqrt{\frac{1}{n} + \frac{1}{m}}} \sim t_{n-1+m-1}$$

   where

$$s^2_p = \frac{(n-1)s^2_X + (m-1)s^2_Y}{n+m-2}$$

   and we can use this to get a confidence interval for $\mu_X - \mu_Y$.

4. From (a) and (e) under the assumptions on the X's and Y's in (g), $\frac{s^2_X/\sigma^2_X}{s^2_Y/\sigma^2_Y} \sim F_{n-1,m-1}$ and we can use this to get a confidence interval for σ^2_Y/σ^2_X.