Sample Mean & Sample Variance

Let $X_1, \ldots, X_n$ be a random sample from $N\left(\mu, \sigma^2\right)$ and let $\bar{X}$ and S^2 be the sample mean and sample variance. Then

- $\bar{X}$ and S^2 are independent random variables. [Hint: Show $\bar{Z}$ and S_Z^2 are independent, each $Z_i = \frac{X_i - \mu}{\sigma} \sim N(0,1)$ and consider the transformation $Y_1 = \bar{Z}$, $Y_i = Z_i - \bar{Z}$ where $i = 2, \ldots, n$.]

- $(n-1)S^2/\sigma^2$ has the chi-square distribution with $n-1$ degrees of freedom.

  [Hint: consider $(n-1)S^2 + n(\bar{X}-\mu)^2 = \sum_{i=1}^n (X_i - \mu)^2$ and use mgf.]

Gamma Distribution

- Suppose $U \sim U(0,1)$, $\beta > 0$. Then $-\beta \ln(U) \sim \mathit{Exp}(\beta) = \mathit{Gamma}(1, \beta)$.

- Suppose all Z_i are independent standard normal random variables, then $\sum_{i=1}^n Z_i^2 \sim \chi^2(k) = \mathit{Gamma}(\frac{n}{2}, 2)$.

- If X_i has a $\mathit{Gamma}(\alpha_i, \beta)$ distribution for $i = 1, \ldots, m$ and all X_i are independent, then $\sum_{i=1}^m X_i \sim \mathit{Gamma}(\sum_{i=1}^m \alpha_i, \beta)$.

- If $X \sim \mathit{Gamma}(\alpha, \beta)$, then for any $c > 0$, $cX \sim \mathit{Gamma}(\alpha, c\beta)$.

- If $X \sim \mathit{Gamma}(\alpha, \beta)$, then

$$E(X^k) = \frac{\Gamma(\alpha + k)}{\Gamma(\alpha)} \beta^k, \ \textit{where } \alpha > -k.$$

Exercise 7.30

Suppose that Z has a standard normal distribution and that Y is an independent χ^2-distributed random variable with n df. Then, according to Definition 7.2,

$$T = \frac{Z}{\sqrt{Y/n}}$$

has a t distribution with n df.

a Give $E(Z)$ and $E\left(Z^2\right)$. [Hint: $E\left(Z^2\right) = V(Z) + (E(Z))^2$.]

b Show that $E(T) = 0$, if $n > 1$.

c Show that $V(T) = n/(n-2)$, if $n > 2$.

Exercise 6.34

Suppose that W_1 and W_2 are independent χ^2-distributed random variables with v_1 and v_2 d.f., respectively. According to Definition 7.3,

$$F = \frac{W_1/v_1}{W_2/v_2}$$

has an F distribution with n and m numerator and denominator degrees of freedom, respectively. Use the preceding structure of F, the independence of W_1 and W_2, and the result summarized in Exercise 7.30(b) to show

a $E(F) = v_2/\left(v_2 - 2\right)$, if $v_2 > 2$.

b $V(F) = \left[2v_2^2\left(v_1 + v_2 - 2\right)\right] / \left[v_1\left(v_2 - 2\right)^2\left(v_2 - 4\right)\right]$, if $v_2 > 4$.

Exercise 7.39

Suppose that independent samples (of sizes n_i) are taken from each of k populations and that population i is normally distributed with mean μ_i and variance $\sigma^2, i = 1, 2, \ldots, k$. That is, all populations are normally distributed with the same variance but with (possibly) different means. Let $\bar{X}_i$ and $S_i^2, i = 1, 2, \ldots, k$ be the respective sample means and variances. Let $\theta = c_1\mu_1 + c_2\mu_2 + \cdots + c_k\mu_k$, where $c_1, c_2, \ldots, c_k$ are given constants.

a Give the distribution of $\hat{\theta} = c_1\bar{X}_1 + c_2\bar{X}_2 + \cdots + c_k\bar{X}_k$.

b Give the distribution of

$$\frac{\text{SSE}}{\sigma^2}, \quad \text{where SSE} = \sum_{i=1}^{k} (n_i - 1) S_i^2.$$

c Give the distribution of

$$\frac{\hat{\theta} - \theta}{\sqrt{\left(\frac{c_1^2}{n_1} + \frac{c_2^2}{n_2} + \cdots + \frac{c_k^2}{n_k}\right) \text{MSE}}}, \quad \text{where MSE } = \frac{\text{SSE}}{n_1 + n_2 + \cdots + n_k - k}.$$

Exercise 6.30

a $E(Z)=0, E\left(Z^2\right)=V(Z)+[E(Z)]^2=1$.

b From the expression of $E(X^k)$ where $X \sim \mathit{Gamma}(\alpha, \beta)$, we have $E(T)=E(Z)E(\frac{1}{\sqrt{Y/n}})=0$ if $n>1$.

c From the expression of $E(X^k)$ where $X \sim \mathit{Gamma}(\alpha, \beta)$, we have

$$V(T)=E\left(T^2\right)=nE\left(Z^2/Y\right)=nE(Z^2)E(Y^{-1})=n/(n-2), n>2.$$

Exercise 6.34

Similar to 6.30, we have:

a $E(F)=\frac{v_2}{v_1}E\left(W_1\right)E\left(W_2^{-1}\right)=\frac{v_2}{v_1}\times\left(\frac{v_1}{v_2-2}\right)=v_2/\left(v_2-2\right), v_2>2$.

b Again from From the expression of $E(X^k)$ where $X \sim \mathit{Gamma}(\alpha, \beta)$, we have

$$\begin{aligned} V(F)=E\left(F^2\right)-[E(F)]^2 &= \left(\frac{v_2}{v_1}\right)^2 E\left(W_1^2\right)E\left(W_2^{-2}\right)-\left(\frac{v_2}{v_2-2}\right)^2 \\ &= \left(\frac{v_2}{v_1}\right)^2 v_1\left(v_1+2\right)\frac{1}{\left(v_2-2\right)\left(v_2-4\right)}-\left(\frac{v_2}{v_2-2}\right)^2 \\ &= \left[2v_2^2\left(v_1+v_2-2\right)\right]/\left[v_1\left(v_2-2\right)^2\left(v_2-4\right)\right], v_2>4. \end{aligned}$$

Exercise 6.39

a Note that for $i=1,2,\ldots,k$, the $\bar{X}_i$ have independent a normal distributions with mean μ_i and variance σ/n_i. Since $\hat{\theta}$, a linear combination of independent normal random variables, by Theorem 6.3, $\hat{\theta}$ has a normal distribution with mean given by

$$E(\hat{\theta})=E\left(c_1\bar{X}_1+\ldots+c_k\bar{X}_k\right)=\sum_{i=1}^{k}c_i\mu_i$$

and variance

$$V(\hat{\theta})=V\left(c_1\bar{X}_1+\ldots+c_k\bar{X}_k\right)=\sigma^2\sum_{i=1}^{k}c_i^2/n_i^2.$$

b For $i=1,2,\ldots,k,\left(n_i-1\right)S_i^2/\sigma^2$ follows a chi-square distribution with n_i-1 degrees of freedom. In addition, since the S_i^2 are independent,

$$\frac{\text{SSE}}{\sigma^2}=\sum_{i=1}^{k}\left(n_i-1\right)S_i^2/\sigma^2$$

is a sum of independent chi-square variables. Thus, the above quantity is also distributed as chi-square with degrees of freedom $\sum_{i=1}^{k}\left(n_i-1\right)=\sum_{i=1}^{k}n_i-k$.

c From part a, we have that $\frac{\hat{\theta}-\theta}{\sigma\sqrt{\sum_{i=1}^{k}c_i^2/n_i^2}}$ has a standard normal distribution. Therefore, by Definition 7.2, a random variable constructed as

$$\frac{\hat{\theta}-\theta}{\sigma\sqrt{\sum_{i=1}^{k}c_i^2/n_i^2}}\Bigg/\sqrt{\frac{\sum_{i=1}^{k}\left(n_i-1\right)S_i^2/\sigma^2}{\sum_{i=1}^{k}n_i-k}}=\frac{\hat{\theta}-\theta}{\sqrt{\text{MSE}\sum_{i=1}^{k}c_i^2/n_i^2}}$$

has the t-distribution with $\sum_{i=1}^{k}n_i-k$ degrees of freedom. Here, we are assuming that $\hat{\theta}$ and SSE are independent (similar to $\bar{Y}$ and S^2 as in Theorem 7.3).