- **1.** MGB VI.18[c,d] On the F distribution:
 - (a) If X has an F distribution with m and n degrees of freedom, show that

$$W = \frac{mX/n}{1 + mX/n}$$

has a beta distribution.

Solution:

Note that except for $W \neq 0$, W is a transformation of X by $g:(0,\infty) \to (0,1)$ where

$$g(x) = \frac{mx/n}{1 + mx/n}.$$

Moreover, the inverse of g is given by

$$g^{-1}(w) = \frac{n}{m} \cdot \frac{w}{1 - w}.$$

so the transformation is one-to-one, and the transformation theorem applies. The Jacobian is

$$\frac{n}{m} \cdot (1-w)^{-2},$$

so W has density

$$f_W(w) = \frac{n}{m} (1 - w)^{-2} f_{F(m,n)}(g^{-1}(w))$$

$$= \frac{n}{m} (1 - w)^{-2} \frac{\left(\frac{m}{n}\right)^{m/2}}{B(m/2, n/2)} \left(\frac{m}{n} \frac{w}{1 - w}\right)^{m/2 - 1} \left(1 + \frac{m \cancel{w}}{\cancel{m}} \frac{w}{1 - w}\right)^{-(m+n)/2} I_{(0,1)}(w)$$

$$= \frac{1}{B(m/2, n/2)} \left(\frac{m}{n}\right)^{-1 + m/2 - m/2 + 1} w^{m/2 - 1} (1 - w)^{-2 + (-m/2 + 1) + (m+n)/2}$$

since
$$1 + w/(1 - w) = 1/(1 - w)$$
,

$$= \frac{1}{B(m/2, n/2)} w^{m/2-1} (1 - w)^{n/2-1}.$$

This is the p.d.f. of a beta distribution with parameters a = m/2 and b = n/2.

(b) Use the result of part (a) and the beta function to find the mean and variance of the F distribution. [Find the first two moments of mX/n = W/(1-W)].

Solution:

The rth moment of X is given by

$$E[X^r] = \left(\frac{n}{m}\right)^r E[(mX/n)^r] = \left(\frac{n}{m}\right)^r E\left[\frac{W^r}{(1-W)^r}\right]$$

$$= \left(\frac{n}{m}\right)^r \frac{1}{B(m/2, n/2)} \int_0^1 w^{m/2-1+r} (1-w)^{n/2-1-r}$$

$$= \left(\frac{n}{m}\right)^r \frac{B(m/2+r, n/2-r)}{(m/2), n/2)}$$

$$= \left(\frac{n}{m}\right)^r \frac{\Gamma(m/2+r)\Gamma(n/2-r)}{\Gamma(m/2)\Gamma(n/2)}.$$

So

$$\mu = \frac{n}{m} \cdot \frac{m/2}{n/2 - 1} = \frac{n}{n - 2},$$

and

$$\sigma^{2} = \left(\frac{n}{m}\right)^{2} \frac{(m/2+1) \cdot m/2}{(n/2-1)(n/2-2)} - \left(\frac{n}{n-2}\right)^{2}$$

$$= \frac{n^{2}(m+2)}{m(n-2)(n-4)} - \frac{n^{2}}{(n-2)^{2}}$$

$$= \frac{n^{2}(m+2)(n-2) - n^{2}m(n-4)}{m(n-2)^{2}(n-4)}$$

$$= \frac{2n^{2} \cdot (m+n-4)}{m(n-2)^{2}(n-4)}.$$

2. MGB VI.19[c,d] On the t distribution:

(a) If X is t-distributed, show that X^2 is F-distributed.

Solution:

Denote $Y = X^2$ and let $g: A_1 \cup A_2 \to (0, \infty)$ by $g(x) = x^2$ where $A_1 = (-\infty, 0)$, and $A_2 = (0, \infty)$. Note that g restricted to each A_i is one-to-one. Therefore, we can apply the transformation theorem:

$$f_Y(y) = \left[\frac{1}{2} \frac{1}{\sqrt{y}} f_X(-\sqrt{y}) + \frac{1}{2} \frac{1}{\sqrt{y}} f_X(\sqrt{y}) \right]$$

$$= \frac{1}{\sqrt{y}} f_X(\sqrt{y}) \qquad \text{since } f_X \text{ is even}$$

$$= \frac{1}{\sqrt{y}} \frac{\Gamma[(k+1)/2]}{\Gamma(k/2)} \frac{1}{\sqrt{k\pi}} \frac{1}{(1+y/k)^{(k+1)/2}}$$

$$= \frac{\Gamma[(1+k)/2]}{\Gamma(1/2)\Gamma(k/2)} \left(\frac{1}{k} \right)^{1/2} \frac{y^{(1-2)/2}}{(1+\frac{1}{k}y)^{(k+1)/2}}.$$

This is the density for a random variable distributed F(1, k).

(b) If X is t-distributed with k degrees of freedom, show that $1/(1 + X^2/k)$ has a beta distribution.

Solution:

Using part (a), this is a corollary of MGB VI.18[d].

3. MGB VI.22 Let $X_1, ..., X_n$ be a random sample from $N(\mu, \sigma^2)$. Define

$$\overline{X}_{k} = \frac{1}{k} \sum_{i=1}^{k} X_{i}, \qquad \qquad S_{k}^{2} = \frac{1}{k-1} \sum_{i=1}^{k} (X_{i} - \overline{X}_{k})^{2},$$

$$\overline{X}_{n-k} = \frac{1}{n-k} \sum_{i=k+1}^{n} X_{i}, \qquad S_{n-k}^{2} = \frac{1}{n-k-1} \sum_{i=k+1}^{n} (X_{i} - \overline{X}_{n-k})^{2},$$

$$\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_{i}, \quad \text{and} \qquad S^{2} = \frac{1}{n-1} \sum_{i=1}^{n} (X_{i} - \overline{X})^{2}.$$

Use known results about sampling from the normal distribution to answer the following:

(a) What is the distribution of $\sigma^{-2}[(k-1)S_k^2 + (n-k-1)S_{n-k}^2]$?

Distributing σ^{-2} , we have that this is the sum of independent Chi-squared distributions with degrees of freedom k-1 and n-k-1 respectively. By independence (moment generating function argument), the sum is distributed χ^2_{n-2} .

(b) What is the distribution of $(\frac{1}{2})(\overline{X}_k + \overline{X}_{n-k})$?

We have that $\overline{X}_k \sim N(\mu, \sigma^2/k)$ and $\overline{X}_{n-k} \sim N(\mu, \frac{\sigma^2}{n-k})$. Hence, by indpendence, their sum is distributed $N\left(2\mu, \frac{\sigma^2n}{k(n-k)}\right)$ since $\frac{1}{k} + \frac{1}{n-k} = \frac{n}{k(n-k)}$. Finally, multiplying by 1/2 yields a distribution of $N\left(\mu, \frac{\sigma^2}{4}(\frac{\sigma^2n}{k(n-k)})\right)$.

- (c) What is the distribution of $\sigma^{-2}(X_i \mu)$? Recall that $Z = (X_i - \mu)/\sigma \sim N(0, 1)$, hence $\sigma^{-2}(X_i - \mu) = \frac{1}{2}Z \sim N(0, \frac{1}{2})$.
 - (d) What is the distribution of S_k^2/S_{n-k}^2 ?

Note that $(k-1)S_k^2/\sigma^2 \sim \chi_{k-1}^2$ and $(n-k-1)S_{n-k}^2/\sigma^2 \sim \chi_{n-k-1}^2$. So

$$\frac{k-1}{n-k-1} \cdot \frac{S_k^2}{S_{n-k}^2} \sim F(k-1, n-k-1).$$

The distribution of S_k^2/S_{n-k}^2 can be obtained by transforming the above F distributed variable by multiplying by (n-k-1)/(k-1). This gives a density

$$(k-1)/(n-k-1)f_Y\left(\frac{k-1}{n-k-1}\right)$$
,

where $Y \sim F(k - 1, n - k - 1)$.

(e) What is the distribution of $(\overline{X} - \mu)/(8/\sqrt{n})$?

If $Z = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}}$, and $Y = \frac{n-1}{\sigma^2} S^2$, then $X \sim N(0, 1)$, $Y \sim \chi_{n-1}^2$ and

$$(\overline{X} - \mu)/(8/\sqrt{n}) = (\sigma/\sqrt{n})Z / \sqrt{\frac{Y\sigma^2}{n(n-1)}} = \sqrt{n-1}W,$$

where $W \sim t(n-1)$. Hence, the desired density is $(n-1)^{-1/2} f_t(w(n-1)^{-1/2})$, where f_t is the density of Student's t distribution.

- **4.** MGB VI.23 Let Z_1, Z_2 be a random sample of size 2 from N(0,1) and X_1, X_2 be a random sample of size 2 from N(1,1). Suppose the $Z_i's$ are independent of the $X_j's$. Use known results about sampling from the normal distribution to answer the following:
 - (a) What is the distribution of $\overline{X} \overline{Z}$?
 - **(b)** What is the distribution of $(Z_1 + Z_2)/\sqrt{[X_2 X_1)^2 + (Z_2 Z_1)^2]/2}$?
 - (c) What is the distribution of $[(X_1 X_2)^2 + (Z_1 Z_2)^2 + (Z_1 + Z_2)^2]/2$?
 - (d) What is the distribution of $(X_2 + X_1 2)^2/(X_2 X_1)^2$?

5. MGB VI.27 If $X_1, X_2, ..., X_n$ are indepently and normally distributed with the same mean but different variances $\sigma_1^2, \sigma_2^2, ..., \sigma_n^2$ and assuming that

$$U = \frac{\sum_{i=1}^{n} X_i / \sigma_i^2}{\sum_{i=j}^{n} 1 / \sigma_j^2}$$
 and $V = \sum_{i=1}^{n} \frac{(X_i - U)^2}{\sigma_i^2}$

are independently distributed, show that U is normal and V has the chi-square distribution with n-1 degrees of feedom.

Solution:

Note that U is a linear combination of normal random variables, hence its distribution is given by

$$U \sim N \left(\mu \frac{\sum_{i=1}^{n} 1/\sigma_i^2}{\sum_{j=1}^{n} 1/\sigma_j^2}, \left[\sum_{i=1}^{n} \frac{\sigma_i^2}{\sigma_i^4} \middle/ \left(\sum_{j=1}^{n} 1/\sigma_j^2 \right)^2 \right] \right)$$
$$= N \left(\mu, \left(\sum_{j=1}^{n} 1/\sigma_j^2 \right)^{-1} \right).$$

To evaluate the distribution of V, note

$$\begin{split} \sum_{i=1}^{n} \frac{(X_i - \mu)^2}{\sigma_i^2} &= \sum_{i=1}^{n} \frac{(X_i - U + U - \mu)^2}{\sigma_i^2} \\ &= \sum_{i=1}^{n} \frac{(X_i - U)^2}{\sigma_i^2} + 2(U - \mu) \sum_{i=1}^{n} \frac{(X_i - U)}{\sigma_i^2} + (U - \mu)^2 \sum_{i=1}^{n} \frac{1}{\sigma_i^2}. \\ &= \sum_{i=1}^{n} \frac{(X_i - U)^2}{\sigma_i^2} + 2(U - \mu) \left[\sum_{i=1}^{n} \frac{X_i}{\sigma_i^2} \cdot U \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \right] + (U - \mu)^2 \sum_{i=1}^{n} \frac{1}{\sigma_i^2}. \end{split}$$

since $U \sum 1/\sigma_i^2 = \sum X_i/\sigma_i^2$

$$\stackrel{\dagger}{=} V + (U - \mu)^2 \sum_{i=1}^n \frac{1}{\sigma_i^2}.$$

We know that $(X_i - \mu)/\sigma_i^2$ is a standard normal random variable so $\sum_{i=1}^n (X_i - \mu)/\sigma_i^2 \sim \chi_n^2$. Moreover, by the above calculation $(U - \mu)/\sqrt{\sum 1/\sigma_i}$ is also a standard normal random variable. Hence, its square (which is the second term on the right hand side of †) is distributed χ_1^2 . The equality of the random variables above implies that their moment generating functions are equal. Moreover, since V is assumed independent of U, it is independent of $(U - \mu)^2 \sum 1/\sigma_i$. Thus the moment generating function of the right hand side of † factors. I.e.

$$\left(\frac{1}{1-2t}\right)^{n/2} = M_V(t) \cdot \left(\frac{1}{1-2t}\right)^{1/2} \iff M_V(t) = \left(\frac{1}{1-2t}\right)^{(n-1)/2}.$$

This is the moment generating function for a χ_{n-1}^2 distributed random variable.

6. MGB VI.29 Let a sample of size n_1 from a normal population (with variance σ_1^2) have sample variance S_1^2 , and let a second sample of size n_2 from a second normal population (with mean μ_2 and variance σ_2^2) have sample mean \overline{X} and sample variance S_2^2 . Find the joint density of

$$U = \frac{\sqrt{n_2}(\overline{X} - \mu_2)}{S_2} \quad \text{and} \quad V = \frac{S_1^2}{S_2^2}$$

(Assume that samples are independent.)

Solution:

Let us first establish a lemma.

Lemma 1. If $Y \sim \chi_k^2$, then $aY \sim Gamma(\frac{1}{2}, \frac{k}{2a})$.

Proof. Consider the moment generating function of aY

$$\begin{split} M_{aY}(t) &= M_Y(at) \\ &= \left(\frac{\frac{k}{2}}{\frac{k}{2} - at}\right)^{\frac{1}{2}} \\ &= \left(\frac{\frac{k}{2a}}{\frac{k}{2a} - t}\right)^{\frac{1}{2}} \\ &= M_W(t) \quad \text{where } W \sim Gamma\left(\frac{1}{2}, \frac{k}{2a}\right). \end{split}$$

Suppose $S_2^2 = s$ is given. Then

$$U|_{S_2^2=s} \sim N(0, \sigma_2^2/s^2)$$
 and $\frac{s^2(n_1-1)}{\sigma_1^2}$

7. Supplement 1. Let $Z_1, Z_2, ...$ be a sequence of random variables; and suppose that, for n = 1, 2, ..., the distribution of Z_n is as follows:

$$P(Z_n = n^2) = \frac{1}{n}$$
 and $P(Z_n = 0) = 1 - \frac{1}{n}$.

Show that

$$\lim_{n\to\infty} E(Z_n) = \infty \text{ but } Z_n \stackrel{p}{\longrightarrow} 0.$$

Solution:

Note that for each n, the events $Z_n = n^2$ and $Z_n = 0$ are disjoint and that the probability of their sum is 1, hence the support of each Z_n is $\{0, n^2\}$ and has the p.m.f.

$$f_{Z_n}(m) = \begin{cases} \frac{1}{n} & \text{if } m = n^2\\ 1 - \frac{1}{n} & \text{if } m = 0\\ 0 & \text{otherwise.} \end{cases}$$

Thus, we can calculate the expected value of each \mathbb{Z}_n as

$$E(Z_n) = \sum_{0,n^2} x f_{Z_n}(x)$$
$$= 0 \left(1 - \frac{1}{n}\right) + n^2 \left(\frac{1}{n}\right)$$
$$= n$$

Hence $E(Z_n) \to \infty$ exactly as $n \to \infty$.

However, suppose that $\varepsilon > 0$ is given. Then

$$\lim_{n \to \infty} P(|Z_n - 0| < \varepsilon) = \lim_{n \to \infty} P(|Z_n| < \varepsilon)$$

$$= \lim_{n \to \infty} P(Z_n = 0) \quad \text{for if } n \ge \sqrt{\varepsilon} \text{ then } 0 < \varepsilon \le n^2,$$

$$= \lim_{n \to \infty} \left(1 - \frac{1}{n}\right)$$

$$= 1.$$

Hence $Z_n \stackrel{p}{\longrightarrow} 0$.

8. Supplement 2. A sequence of random variables $Y_1, Y_2, ...$ is said to converge in the rth mean if $\lim_{n\to\infty} E(|Y_n-b|^r)=0$. Prove that if a sequence of random variables converge to b in the quadratic mean, then the sequence also converges to b in probability.

Solution:

Let $\varepsilon > 0$ be given. First note that because the function $f(x) = x^2$ is increasing for x > 0,

$$P[|Y_n - b| < \varepsilon] = P[(Y_n - b)^2 < \varepsilon^2]$$

$$\stackrel{\dagger}{=} 1 - P[(Y_n - b)^2 \ge \varepsilon^2].$$

Recall Markov's inequality (denoted Chebyechev's Inequality in MGB) guarantees,

$$P[(Y_n - b)^2 \ge \varepsilon^2] \le \frac{E[(Y_n - b)^2]}{\varepsilon^2},$$

Hence,

$$1 - \frac{E\left[(Y_n - b)^2\right]}{\varepsilon^2} \le 1 - P\left[(Y_n - b)^2 \ge \varepsilon^2\right]$$

$$\iff 1 - \frac{E\left[(Y_n - b)^2\right]}{\varepsilon^2} \stackrel{\dagger}{\le} P\left[|Y_n - b| < \varepsilon\right]$$

$$\iff \lim_{n \to \infty} \left\{1 - \frac{E\left[(Y_n - b)^2\right]}{\varepsilon^2}\right\} \le \lim_{n \to \infty} P\left[|Y_n - b| < \varepsilon\right]$$

$$\iff 1 - 0 \le \lim_{n \to \infty} P\left[|Y_n - b| < \varepsilon\right].$$

Since probability is bounded above by 1, we have $\lim_{n\to\infty} P[|Y_n - b| < \varepsilon] = 1$ and, thus, convergence in probability.

9. Supplement 3. Let $X_1, X_2, ...$ be a sequence of random variables. By the Weak Law of Large Numbers (provided that $E(X^4) < \infty$) we have

$$\overline{X}_n \stackrel{p}{\longrightarrow} \mu$$
 and $\frac{1}{n} \sum_{i=1}^n X_i^2 \stackrel{p}{\longrightarrow} E(X^2)$.

Use these to prove that $\widehat{\sigma}_n^2 \stackrel{p}{\longrightarrow} \sigma^2$ where the sample variance $\widehat{\sigma}_n^2$ is defined by

$$\widehat{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \overline{X}_n)^2.$$

(Hint: Define $g: \mathbb{R}^2 \to \mathbb{R}$ as $g(y, z) = y - z^2$ which is a continuous function.) Solution:

Consider the joint random variable $(\frac{1}{n}\sum_{i=1}^{n}X_{i}^{2},\overline{X}_{n})$ and the continuous transformation $g:\mathbb{R}^{2}\to\mathbb{R},\ g(y,z)=y-z^{2}$. Then, by the statements above and Theorem*,

$$g\left(\frac{1}{n}\sum_{i=1}^{n}X_{i}^{2},\overline{X}_{n}\right) \stackrel{p}{\longrightarrow} g\left(E(X^{2}),\mu\right) = E(X^{2}) - \mu^{2} = \sigma^{2}.$$

We now calculate,

$$g\left(\frac{1}{n}\sum_{i=1}^{n}X_{i}^{2}, \overline{X}_{n}\right) = \frac{1}{n}\sum_{i=1}^{n}X_{i}^{2} - \overline{X}_{n}^{2}$$

$$= \frac{1}{n}\sum_{i=1}^{n}X_{1}^{2} - 2\overline{X}_{n}^{2} + \overline{X}_{n}^{2}$$

$$= \frac{1}{n}\sum_{i=1}^{n}X_{1}^{2} - 2\left(\frac{1}{n}\sum_{i=1}^{n}X_{i}\right)\overline{X}_{n} + \overline{X}_{n}^{2}$$

$$= \frac{1}{n}\sum_{i=1}^{n}(X_{i} - \overline{X}_{n})^{2}.$$

Theorem*: Suppose $\mathbf{Y_n}: \Omega \to \mathbb{R}^n$ is a random variable such that $\mathbf{Y_n} \stackrel{p}{\longrightarrow} \mathbf{b}$. Then, if $g: \mathbb{R}^n \to \mathbb{R}^m$ is a continuous transformation then $g(\mathbf{Y_n}) \stackrel{p}{\longrightarrow} g(\mathbf{b})$.

Proof. Let $\varepsilon > 0$ and $0 < \delta < 1$. By continuity of g, there exists $\widehat{\delta} > 0$ such that if $|\boldsymbol{y} - \boldsymbol{b}| < \widehat{\delta}$ then $|g(\boldsymbol{y}) - g(\boldsymbol{b})| < \varepsilon$. In the context of the random variable Y_n , $\{e \in \Omega : |\boldsymbol{Y}_n(e) - \boldsymbol{b}| < \widehat{\delta}\} \subseteq \{e \in \Omega : |g(\boldsymbol{Y}_n(e)) - g(\boldsymbol{b})| < \varepsilon\}$. By hypothesis, we can choose N so that if $n \geq N$ implies $P(|\boldsymbol{Y}_n - \boldsymbol{b}| < \widehat{\delta}) \geq 1 - \delta$. Note that because of the containment[†] mentioned, $P(\{e \in \Omega : |g(\boldsymbol{Y}_n(e)) - g(\boldsymbol{b})| < \varepsilon\}) \geq P(\{e \in \Omega : |\boldsymbol{Y}_n(e) - \boldsymbol{b}| < \widehat{\delta})\}) \geq 1 - \delta$, and, thus we have satisfied the desired convergence in probability.

10. Supplement 4. Suppose that X_1, \ldots, X_n form a random sample from a normal distribution with mean 0 and unknown variance σ^2 .

(a) Determine the asymptotic distribution of the statistic $\left(\frac{1}{n}\sum_{i=1}^{n}X_{i}^{2}\right)^{-1}$. Solution:

Recall that the square of a standard normal random variable is distributed chisquared with one degree of freedom. Hence, $X_i^2/\sigma \sim \chi_1^2$. Thus, X_i^2 is a linear transformation of a chi-squared random variable with one degree of freedom. I.e. $X_i^2 = \sigma^2 Y_i^2$, where $Y_i \sim \chi_1^2$. This is a gamma distribution with $\lambda = 1/(2\sigma^2)$ and r = 1/2. To see this, consider its moment generating function Thus, we can think of $X_1^2, X_2^2, ..., X_n^2$ as a random sample from gamma distribution with $\lambda = 1/(2\sigma^2)$ and r = 1/2. Denote them as W_i and their sample mean as \overline{W} . The mean and variance of this gamma distribution are given by $\mu_W = r/\lambda = \sigma^2$ and $\sigma_W^2 = r/\lambda^2 = 2\sigma^4$, respectively. If $g(w) = w^{-1}$, then note that $g'(\mu_W) = -1/(\sigma^4) \neq 0$. By the delta method theorem we have that

$$\frac{\sqrt{n}\left[(\overline{W})^{-1} - (\sigma^2)^{-1}\right]}{(-1/\sigma^4)\sqrt{2\sigma^4}} \stackrel{D}{\longrightarrow} Z \sim N(0,1).$$

Hence $\left(\frac{1}{n}\sum_{i=1}^{n}X_{i}^{2}\right)^{-1}=(\overline{W})^{-1}$ is asymptotically distributed $N(\sigma^{-2},\frac{2}{n\sigma^{4}})$ (removing the negative by symmetry of the normal).

(b) Find a variance stabilizing transformation for the statistic $\frac{1}{n} \sum_{i=1}^{n} X_i^2$. Solution:

Note that $\sigma_W = \sqrt{2}\sigma^2 = \sqrt{2}\mu_W$. To stabilize the variance, we seek g such that

$$1 = g'(\mu_W)\sigma_w \iff g(\mu_W) = \int \sigma_w^{-1} d\mu_W = \frac{1}{\sqrt{2}} \int \mu_W^{-1} d\mu_W$$
$$= \frac{1}{\sqrt{2}} (\ln|\mu_W| + c).$$

Hence, if we transform the data by $g(W) = \ln(W)$, we can expect that the variance the of the transformed asymptotic distribution will not depend on the mean.