1. Sampling Distributions

χ_r^2 **distribution:** A special case of the gamma distribution in which $\gamma = r/2$ (where r is a positive integer) and $\lambda = 1/2$ with pdf as follows:

$$f(x) = \begin{cases} \frac{1}{\Gamma(r/2)2^{r/2}} x^{r/2-1} e^{-x/2} & x > 0 \\ 0 & \text{elsewhere,} \end{cases}$$

is a chi-square distribution with r degrees of freedom (df).

Exercise 1. Let $Z_1, \ldots, Z_n$ be i.i.d. $N(0,1)$ random variables. Then, $Z_1^2 + \cdots + Z_n^2 \sim \text{Gamma}(n/2, 1/2)$.
Solution: For $t > 0$, we have

$$\mathbf{P}\{Z_1^2 \le t\} = \mathbf{P}\{-\sqrt{t} \le Z_1 \le \sqrt{t}\} = 2 \int_0^{\sqrt{t}} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du = \frac{1}{\sqrt{2\pi}} \int_0^t e^{-s/2} s^{-1/2} ds.$$

Differentiate w.r.t t to see that the density of Z_1^2 is $h(t) = \frac{1}{\sqrt{2\pi}} e^{-t/2} t^{-1/2}$, which is just the Gamma$(\frac{1}{2}, \frac{1}{2})$ density.

Now, each Z_k^2 has the same Gamma$(\frac{1}{2}, \frac{1}{2})$ density, and they are independent. Check that when we add independent Gamma random variables with the same scale parameter, the sum has a Gamma distribution with the same scale, but whose shape parameter is the sum of the shape parameters of the individual summands. Therefore, $Z_1^2 + \cdots + Z_n^2$ has Gamma$(n/2, 1/2) \equiv \chi_n^2$ distribution. This completes the solution to the exercise.

t-distribution: Let W denote a random variable that is $N(0,1)$ and V denote a random variable that is χ_r^2, with W and V independent. Then, the joint pdf of W and V (say, $h(w,v)$) is

$$h(w,v) = \begin{cases} \frac{1}{\sqrt{2\pi}} e^{-w^2/2} \frac{1}{\Gamma(r/2)2^{r/2}} v^{r/2-1} e^{-v/2} & -\infty < w < \infty, \quad 0 < v < \infty \\ 0 & \text{elsewhere.} \end{cases}$$

Define a new random variable T by writing

$$T = \frac{W}{\sqrt{V/r}}.$$

The change of variable technique is used to obtain the pdf g_1 of T. The equations

$$t = \frac{w}{\sqrt{v/r}} \quad \text{and} \quad u = v$$

define a one-to-one and onto transformation. Since $w = t\sqrt{u}/\sqrt{r}$ and $v = u$, the absolute value of the Jacobian of the transformation is $|J| = \sqrt{u}/\sqrt{r}$. Accordingly, the joint pdf of T and $U = V$ is given by $g(t,u) = h\left(\frac{t\sqrt{u}}{\sqrt{r}}, u\right)|J|$. The marginal pdf of T is then

$$\begin{aligned}
g_1(t) &= \int_{-\infty}^{\infty} g(t,u)du \\
&= \int_0^{\infty} \frac{1}{\sqrt{2\pi r}\Gamma(r/2)2^{r/2}} u^{(r+1)/2-1} \exp\left[-\frac{u}{2}\left(1+\frac{t^2}{r}\right)\right] du.
\end{aligned}$$

Let $z = u\left[1 + \left(t^2/r\right)\right]/2$, and it is seen that

$$\begin{aligned}
g_1(t) &= \int_0^{\infty} \frac{1}{\sqrt{2\pi r}\Gamma(r/2)2^{r/2}} \left(\frac{2z}{1+t^2/r}\right)^{(r+1)/2-1} e^{-z}\left(\frac{2}{1+t^2/r}\right) dz \\
&= \frac{\Gamma[(r+1)/2]}{\sqrt{\pi r}\Gamma(r/2)} \frac{1}{(1+t^2/r)^{(r+1)/2}}, \quad -\infty < t < \infty.
\end{aligned}$$

Thus, if W is $N(0,1)$, V is χ^2_r and W and V are independent, then $T = W/\sqrt{V/r}$ has the pdf g_1. The distribution of the random variable T is usually called a t-distribution. It should be observed that a t-distribution is completely determined by the parameter r, the degrees of freedom (df) of the chi-square distribution.

F**-distribution:** Consider two independent chi-square random variables U and V having r_1 and r_2 degrees of freedom, respectively. The joint pdf $h(u,v)$ of U and V is then

$$h(u,v) = \begin{cases} \dfrac{1}{\Gamma(r_1/2)\Gamma(r_2/2)2^{(r_1+r_2)/2}} u^{r_1/2-1}v^{r_2/2-1}e^{-(u+v)/2} & 0 < u, v < \infty \\ 0 & \text{elsewhere.} \end{cases}$$

We define the new random variable

$$W = \frac{U/r_1}{V/r_2}$$

and propose finding the pdf g_1 of W. The equations

$$w = \frac{u/r_1}{v/r_2}, \quad z = v$$

define a one-to-one transformation. Since $u = (r_1/r_2)zw$ and $v = z$, the absolute value of the Jacobian of the transformation is $(r_1/r_2)z$. The joint density $g(w,z)$ of W and $Z = V$ is then

$$g(w,z) = \frac{1}{\Gamma(r_1/2)\Gamma(r_2/2)2^{(r_1+r_2)/2}} \left(\frac{r_1 zw}{r_2}\right)^{\frac{r_1-2}{2}} z^{\frac{r_2-2}{2}} \exp\left[-\frac{z}{2}\left(\frac{r_1 w}{r_2}+1\right)\right] \frac{r_1 z}{r_2}.$$

The marginal pdf g_1 of W is then

$$\begin{aligned}
g_1(w) &= \int_{-\infty}^{\infty} g(w,z)dz \\
&= \int_0^{\infty} \frac{(r_1/r_2)^{r_1/2}\, w^{r_1/2-1}}{\Gamma(r_1/2)\,\Gamma(r_2/2)\, 2^{(r_1+r_2)/2}} z^{(r_1+r_2)/2-1} \exp\left[-\frac{z}{2}\left(\frac{r_1 w}{r_2}+1\right)\right] dz.
\end{aligned}$$

If we change the variable of integration by writing

$$y = \frac{z}{2}\left(\frac{r_1 w}{r_2}+1\right),$$

it can be seen that

$$\begin{aligned}
g_1(w) &= \int_0^{\infty} \frac{(r_1/r_2)^{r_1/2}\, w^{r_1/2-1}}{\Gamma(r_1/2)\,\Gamma(r_2/2)\, 2^{(r_1+r_2)/2}} \left(\frac{2y}{r_1 w/r_2+1}\right)^{(r_1+r_2)/2-1} e^{-y}\left(\frac{2}{r_1 w/r_2+1}\right) dy \\
&= \begin{cases} \dfrac{\Gamma\left[(r_1+r_2)/2\right](r_1/r_2)^{r_1/2}}{\Gamma(r_1/2)\,\Gamma(r_2/2)} \dfrac{w^{r_1/2-1}}{(1+r_1 w/r_2)^{(r_1+r_2)/2}} & 0 < w < \infty \\ 0 & \text{elsewhere.} \end{cases}
\end{aligned}$$

Accordingly, if U and V are independent chi-square variables with r_1 and r_2 degrees of freedom (df), respectively, then $W = (U/r_1)/(V/r_2)$ has the pdf g_1. The distribution of this random variable is usually called an F-distribution; and we often call the ratio, which we have denoted by W, as F. That is,

$$F = \frac{U/r_1}{V/r_2}.$$

It should be observed that an F-distribution is completely determined by the two parameters r_1 and r_2.

Student's Theorem: Our final theorem in this section concerns an important result for the later sections on inference for normal random variables. It is a corollary to the t-distribution derived above, and is often referred to as Student's Theorem.

Theorem 2. *Let $X_1, \ldots, X_n$ be i.i.d. $N(\mu, \sigma^2)$. Define the random variables*

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \text{ and } S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2.$$

Then,

(a) $\bar{X}$ has a $N\left(\mu, \frac{\sigma^2}{n}\right)$ distribution,

(b) $\bar{X}$ and S^2 are independent,

(c) $(n-1)S^2/\sigma^2$ *has a* χ^2_{n-1} *distribution, and*
(d) the random variable

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

has a Student t *-distribution with* $(n-1)$ *degrees of freedom.*

Proof. Let $\mathbf{X} = (X_1, \ldots, X_n)$. Since $X_1, \ldots, X_n$ are i.i.d. $N\left(\mu, \sigma^2\right)$ random variables, $\mathbf{X}$ has a multivariate normal distribution $N_n\left(\mu 1_n, \sigma^2 I_n\right)$, where 1_n denotes a vector whose components are all 1 and I_n is the identity matrix. Let $\mathbf{v} = (1/n, \ldots, 1/n) = (1/n)1_n$. Note that $\bar{X} = \mathbf{v}'\mathbf{X}$. Define the random vector $\mathbf{Y}$ by $\mathbf{Y} = \left(X_1 - \bar{X}, \ldots, X_n - \bar{X}\right)$. Consider the following transformation

$$\mathbf{W} = \begin{bmatrix} \bar{X} \\ \mathbf{Y} \end{bmatrix} = \begin{bmatrix} \mathbf{v}' \\ I_n - 1_n\mathbf{v}' \end{bmatrix} \mathbf{X}.$$

Since W is a linear transformation of multivariate normal random vector, it has a multivariate normal distribution with mean

$$\mathbf{E}[\mathbf{W}] = \begin{bmatrix} \mathbf{v}' \\ I_n - 1_n\mathbf{v}' \end{bmatrix} \mu 1_n = \begin{bmatrix} \mu \\ 0_n \end{bmatrix},$$

and covariance matrix

$$\mathrm{Var}[\mathbf{W}] = \begin{bmatrix} \mathbf{v}' \\ I_n - 1_n\mathbf{v}' \end{bmatrix} \sigma^2 I_n \begin{bmatrix} \mathbf{v}' \\ I_n - 1_n\mathbf{v}' \end{bmatrix}' = \sigma^2 \begin{bmatrix} \frac{1}{n} & 0_n' \\ 0_n & I_n - 1_n\mathbf{v}' \end{bmatrix}.$$

Because $\bar{X}$ is the first component of $\mathbf{W}$, we obtain part (a).

Next, because the covariances are 0_n, $\bar{X}$ is independent of $\mathbf{Y}$ (How? - recall the definition of $\mathbf{W}$). But, $S^2 = (n-1)^{-1}\mathbf{Y}'\mathbf{Y}$. Hence, $\bar{X}$ is independent of S^2. Thus, part (b) is true.

Consider the random variable

$$V = \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma}\right)^2.$$

Each term in this sum is the square of a $N(0,1)$ random variable, and hence has a χ^2_1 distribution. Because the summands are independent, it follows that V is a χ^2_n random variable. Note the following identity:

$$V = \sum_{i=1}^{n} \left(\frac{(X_i - \bar{X}) + (\bar{X} - \mu)}{\sigma}\right)^2 = \sum_{i=1}^{n} \left(\frac{X_i - \bar{X}}{\sigma}\right)^2 + \left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}\right)^2 = \frac{(n-1)S^2}{\sigma^2} + \left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}\right)^2.$$

By part (b), the two terms on the right side of the last equation are independent. Further, the second term is the square of a standard normal random variable and hence, has a χ^2_1 distribution. Solving for the mgf of $(n-1)S^2/\sigma^2$ on the right side, we obtain part (c) (How?).

Finally, part (d) follows immediately from parts (a)-(c) (Why?) by writing T as

$$T = \frac{(\bar{X} - \mu)/(\sigma/\sqrt{n})}{\sqrt{(n-1)S^2/\left(\sigma^2(n-1)\right)}} = \frac{\sqrt{n}(\bar{X} - \mu)}{S}.$$

■

2. ORDER STATISTICS

Let $X_1, X_2, \ldots, X_n$ denote a random sample (i.i.d.) from a distribution of continuous type with pdf f that has support $\mathcal{S} = (a, b)$, where $-\infty \leq a < b \leq \infty$. Let Y_1 be the smallest of these X_is; Y_2 the next Xs in order of magnitude; ... and Y_n the largest of X_is. That is, $Y_1 < Y_2 < \cdots < Y_n$ represent $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ when the latter are arranged in ascending order of magnitude. We call Y_i the ith order statistic of the random sample $X_1, X_2, \ldots, X_n$ for $i = 1, 2, \ldots, n$.

The joint pdf of $Y_1, Y_2, \ldots, Y_n$ is given by

$$g(y_1, y_2, \ldots, y_n) = \begin{cases} n! f(y_1) f(y_2) \cdots f(y_n) & a < y_1 < y_2 < \cdots < y_n < b \\ 0 & \text{elsewhere.} \end{cases}$$

Note that the support of $X_1, X_2, \ldots, X_n$ can be partitioned into $n!$ mutually disjoint sets that map onto the support of $Y_1, Y_2, \ldots, Y_n$, namely, $\{(y_1, y_2, \ldots, y_n) : a < y_1 < y_2 < \cdots < y_n < b\}$. One of these $n!$ sets is $a < x_1 < x_2 < \cdots < x_n < b$, and the others can be found by permuting the xs in all possible ways. The transformation associated with the one listed is $x_1 = y_1, x_2 = y_2, \ldots, x_n = y_n$, which has a Jacobian equal to 1. However, the Jacobian of each of the other transformations is either of ± 1. Thus, we have

$$\begin{aligned} g(y_1, y_2, \ldots, y_n) &= \sum_{i=1}^{n!} |J_i| f(y_1) f(y_2) \cdots f(y_n) \\ &= \begin{cases} n! f(y_1) f(y_2) \cdots f(y_n) & a < y_1 < y_2 < \cdots < y_n < b \\ 0 & \text{elsewhere.} \end{cases} \end{aligned}$$

Let $X \sim f$. The distribution function F of X may be written as

$$F(x) = \int_a^x f(w) dw, \quad a < x < b.$$

Observe that

$$\int_a^x [F(w)]^{\alpha - 1} f(w) dw = \frac{[F(x)]^\alpha}{\alpha}, \quad \alpha > 0$$

and

$$\int_y^b [1 - F(w)]^{\beta - 1} f(w) dw = \frac{[1 - F(y)]^\beta}{\beta}, \quad \beta > 0.$$

It is easy to express the marginal pdf of any order statistic (say, Y_k) in terms of F and f. This is done by evaluating the integral

$$g_k(y_k) = \int_a^{y_k} \cdots \int_a^{y_2} \int_{y_k}^b \cdots \int_{y_{n-1}}^b n! f(y_1) f(y_2) \cdots f(y_n) \, dy_n \cdots dy_{k+1} dy_1 \cdots dy_{k-1}.$$

The result is

$$g_k(y_k) = \begin{cases} \frac{n!}{(k-1)!(n-k)!}\left[F(y_k)\right]^{k-1}\left[1-F(y_k)\right]^{n-k} f(y_k) & a < y_k < b \\ 0 & \text{elsewhere.} \end{cases}$$

The joint pdf of any two order statistics (say, $Y_i < Y_j$) is also easily expressed in terms of F and f. We have

$$g_{ij}(y_i, y_j) = \int_a^{y_i} \cdots \int_a^{y_2} \int_{y_i}^{y_j} \cdots \int_{y_{j-2}}^{y_j} \int_{y_j}^{b} \cdots \int_{y_{n-1}}^{b} n! f(y_1) \cdots f(y_n)\, dy_n \cdots dy_{j+1} dy_{j-1} \cdots dy_{i+1} dy_1 \cdots dy_{i-1}.$$

Since, for $\gamma > 0$

$$\int_x^y [F(y) - F(w)]^{\gamma - 1} f(w) dw = - \left. \frac{[F(y) - F(w)]^{\gamma}}{\gamma} \right|_x^y = \frac{[F(y) - F(x)]^{\gamma}}{\gamma},$$

it is found that

$$g_{ij}(y_i, y_j) = \begin{cases} \frac{n!}{(i-1)!(j-i-1)!(n-j)!}\left[F(y_i)\right]^{i-1}\left[F(y_j) - F(y_i)\right]^{j-i-1}\left[1-F(y_j)\right]^{n-j} f(y_i) f(y_j) & a < y_i < y_j < b \\ 0 & \text{elsewhere.} \end{cases}$$

Remark 3. (*Heuristic Derivation*) There is an easy method of remembering the pdf of a vector of order statistics such as the one given above. The probability $\mathbf{P}(y_i < Y_i < y_i + \Delta_i, y_j < Y_j < y_j + \Delta_j)$, where Δ_i and Δ_j are small, can be approximated by the following multinomial probability. In n independent trials, $i - 1$ outcomes must be less than y_i [an event that has probability $p_1 = F(y_i)$ on each trial]; $j - i - 1$ outcomes must be between $y_i + \Delta_i$ and y_j [an event with approximate probability $p_2 = F(y_j) - F(y_i)$ on each trial]; $n - j$ outcomes must be greater than $y_j + \Delta_j$ [an event with approximate probability $p_3 = 1 - F(y_j)$ on each trial]; one outcome must be between y_i and $y_i + \Delta_i$ [an event with approximate probability $p_4 = f(y_i)\Delta_i$ on each trial]; and finally, one outcome must be between y_j and $y_j + \Delta_j$ [an event with approximate probabilty $p_5 = f(y_j)\Delta_j$ on each trial]. This multinomial probability is

$$\frac{n!}{(i-1)!(j-i-1)!(n-j)!\,1!\,1!} p_1^{i-1} p_2^{j-i-1} p_3^{n-j} p_4 p_5,$$

which is just $g_{ij}(y_i, y_j)\Delta_i \Delta_j$.