# $t$- and $F$-distributions

## Useful distributions

Francesco Corona

Department of Computer Science
Federal University of Ceará, Fortaleza

# $t$- and $F$-distributions

The $t$-distribution and the $F$-distribution

- $\rightsquigarrow$ Useful in statistical inference

# The $t$-distribution

## Useful distributions

# The $t$-distribution

Let $W$ indicate a random variable that is $N(0,1)$

$$f(w) = \frac{1}{\sqrt{2\pi}} e^{-w^2/2}, \quad -\infty < w < \infty, \text{ zero elsewhere}$$

Let $V$ indicate a random variable that is $\chi^2(r)$

$$f(v) = \frac{1}{\Gamma(r/2)2^r} v^{r/2-1} e^{-x/2}, \quad 0 < v < \infty, \text{ zero elsewhere}$$

# The $t$-distribution

Let $W$ and $V$ be independent

The joint PDF of $W$ and $V$

$$h(v, w) = \underbrace{\frac{1}{\sqrt{2\pi}} e^{-w^2/2}}_{\mathcal{N}(0,1)} \underbrace{\frac{1}{\Gamma(r/2) 2^{r/2}} v^{r/2-1} e^{-v/2}}_{\chi^2(r)}$$

$$-\infty < w < \infty, 0 < v < \infty, \text{ zero elsewhere}$$

# The $t$-distribution (cont.)

We define a new random variable

$$T = \frac{W}{\sqrt{V/r}} \tag{1}$$

The change-of-variable technique can be used to get the PDF of $T$

The transformation equations

- $t = w/\sqrt{(v/r)}$
- $u = v$

The sets

$$\mathcal{S} = \big\{(w, v) : -\infty < w < \infty, 0 < v < \infty\big\}$$
$$\mathcal{T} = \big\{(t, u) : -\infty < t < \infty, 0 < u < \infty\big\}$$

The inverse transformation equation

- $w = t\sqrt{u}/\sqrt{r}$
- $v = u$

The absolute value of the Jacobian of the transformation $|J| = \sqrt{u}/\sqrt{r}$

# The $t$-distribution (cont.)

The joint PDF of $T$ and $U = V$

$$g(t, u) = h\Big(\frac{t\sqrt{u}}{\sqrt{r}}, u\Big) = \frac{1}{\sqrt{2\pi}\Gamma(r/2)2^{r/2}} u^{r/2-1} e^{\left[-\frac{u}{2}\left(1+\frac{t^2}{2}\right)\right]} \frac{\sqrt{u}}{\sqrt{r}}$$

$$|t| < \infty, 0 < u < \infty, \text{ zero elsewhere}$$

# The $t$-distribution (cont.)

The marginal PDF of $T$

$$g_T(t) = \int_{-\infty}^{\infty} g(t,u)\mathrm{d}u$$
$$= \int_0^{\infty} \frac{1}{\sqrt{2\pi r}\Gamma(r/2)2^{r/2}} u^{(r+1)/2-1} e^{\left[-\frac{u}{2}\left(1+\frac{t^2}{r}\right)\right]}\mathrm{d}u$$

In the integral, we let $z = u\left[1 + (t^2/r)\right]/2$

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

# The $t$-distribution (cont.)

$$\frac{\Gamma\big[(r+1)/2\big]}{\sqrt{\pi r}\Gamma(r/2)}\frac{1}{(1+t^2/r)^{(r+1)/2}}, \quad -\infty < t < \infty$$

The distribution of the random variable $T$ is called the ***t*-distribution**

The $t$-distribution is completely determined by parameter $r$

- ⇝ The number of degrees of freedom

Table of approximate values of probability for selected $r$ and $t$

$$P(T \leq t) \int_{-\infty}^{t} g_T(t)\mathrm{d}t$$

As the degrees of freedom goes $\infty$, the $t$-distribution converges to $N(0,1)$

# The $t$-distribution (cont.)

## Example

### Mean and variance of the $t$-distribution

Let the random variable $T$ have the $t$-distribution, $r$ degrees of freedom

We can write $T = W(V/r)^{-1/2}$, with $W \sim N(0,1)$ and $V \sim \chi^2(r)$

- Let $W$ and $V$ be independent RVs

Provided $(r/2) - (k/2) > 0$ (that is, $k < r$), by independence

$$\begin{aligned} E(T^k) &= E\Big[W^k\Big(\frac{V}{r}\Big)^{-k/2}\Big] = E(W^k)E\Big[\Big(\frac{V}{r}\Big)^{-k/2}\Big] \\ &= E(W^k)\frac{2^{-k}\Gamma(r/2-k/2)}{\Gamma(r/2)r^{-k/2}}, \quad \text{if } k < r \end{aligned} \tag{3}$$

■

# The *F*-distribution

## Useful distributions

# The $F$-distribution

Let $U$ and $V$ be two independent random variables with $r_1$ and $r_2$ DOFs

The joint PDF $h(u, v)$ of $U$ and $V$

$$h(u, v) = \frac{1}{\Gamma(r_1/2)\Gamma(r/2)2^{(r_1+r_2)/2}} u^{(r_1/2-1)} v^{(r_2/2-1)} e^{-(u+v)/2}$$
$$0 < u, v < \infty$$

We define the new random variable

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

We are interested in the PDF $g_W(w)$ of $W$

# The $F$-distribution (cont.)

The transformation equations

- $w = (u/r_1)/(v/r_2)$
- $z = v$

The sets

$$\mathcal{S} = \{(u, v) : 0 < u < \infty, 0 < v < \infty\}$$
$$\mathcal{T} = \{(w, z) : 0 < w < \infty, 0 < z < \infty\}$$

The inverse transformation equatons

- $u = (r_1/r_2)zw$
- $v = z$

The absolute value of the Jacobian of the transformation $|J| = (r_1/r_2)z$

# The $F$-distribution (cont.)

The joint PDF of $W$ and $Z = V$

$$g(w, z) = \frac{1}{\Gamma(r_1/2)\Gamma(r_2/2)2^{(r_1+r_2)/2}} \Big(\frac{r_1 z w}{r_2}\Big)^{(\frac{r_1-2}{2})} z^{(\frac{r_2-2}{2})} e^{\Big[-\frac{z}{2}\Big(\frac{r_1 w}{r_2}+1\Big)\Big]} \frac{r_1 z}{r_2}$$

For $(w, z) \in \mathcal{T}$ and zero elsewhere

# The $F$-distribution (cont.)

The marginal PDF of $W$

$$g_W(w) = \int_{-\infty}^{\infty} g(w,z)\mathrm{d}z$$

$$= \int_0^{\infty} \frac{(r_1/r_2)^{(r_1/2)}(w)^{(r_1/2-1)}}{\Gamma(r_1/2)\Gamma(r_2/2)2^{(r_1+r_2)/2}} z^{(r_1+r_2)/2-1} e^{\left[-\frac{z}{2}\left(\frac{r_1 w}{r_2}+1\right)\right]}\mathrm{d}z$$

In the integral, we let $y = z/2\big(r_1 w/r_2 + 1\big)$

$$g_W(w) = \int_0^{\infty} \frac{(r_1/r_2)^{(r_1/2)}(w)^{(r_1/2-1)}}{\Gamma(r_1/2)\Gamma(r_2/2)2^{(r_1+r_2)/2}} \left(\frac{2y}{r_1 w/r_2 + 1}\right)^{(r_1+r_2)/2-1} e^{-y}\left(\frac{2}{r_1 w/r_2+1}\right)\mathrm{d}y \qquad (4)$$

$$= \frac{\Gamma\big[(r_1+r_2)/2\big](r_1/r_2)^{(r_1/2)}}{\Gamma(r_1/2)\Gamma(r_2/2)} \frac{(w)^{r_1/2-1}}{(1+r_1 w/r_2)^{(r_1+r_2)/2}}$$

For $0 < w < \infty$ and zero elsewhere

# The $F$-distribution (cont.)

$$g_W(w) = \frac{\Gamma\big[(r_1 + r_2)/2\big](r_1/r_2)^{(r_1/2)}}{\Gamma(r_1/2)\Gamma(r_2/2)} \frac{(w)^{r_1/2-1}}{(1 + r_1 w/r_2)^{(r_1+r_2)/2}}$$

The distribution of the random variable $W$ (F) is called the **$F$-distribution**

The $F$-distribution is completely determined by two parameters $r_1$ and $r_2$

Table of approximated values of the probability for selected $r_1$, $r_2$ and $b$

$$P(F \leq b) = \int_0^b g_W(w) \mathrm{d}w$$

# The $F$-distribution (cont.)

## Example

**Moments of the $F$-distribution**

Let the random variable $F$ have the $F$-distribution, $r_1$ and $r_2$ DOFs

We can write $F = (r_2/r_1)/(U/V)$, with $U \sim \chi^2(r_1)$ and $V \sim \chi^2(r_2)$

- Let $U$ and $V$ be independent RVs

By independence,

$$E(F^k) = \left(\frac{r_2}{r_1}\right)^k E(U^k)E(V^{-k})$$

Provided that both expectations exist

# The $F$-distribution (cont.)

Because $k > -(r_1/2)$ is always true, the first expectation always exists

The second one, exists if $r_2 > 2k$ (the denominator DOFs must exceed $2k$)

Assuming this is true,

$$E(F) = \frac{r_2}{r_1} r_1 \frac{2^{-1}\Gamma(r_2/2-1)}{\Gamma(r_2/2)} = \frac{r_2}{r_2-2} \qquad (5)$$

■

# Student's theorem

## Useful distributions

# Student's theorem

An important result for inference of normal random variables

- It is a corollary to the $t$-distribution

$\rightsquigarrow$ Student's theorem

# Student's theorem (cont.)

## Theorem

***Student's theorem***

*Let $X_1, \ldots, X_n$ be IID random variables*

*Let each each $X_i \sim N(\mu, \sigma_2$*

*Define the random variables*

$$\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

$$S^2 = \frac{1}{n = 1} \sum_{i=1}^{n} (X_i - \overline{X})^2$$

*Then,*

(a) $\overline{X} \sim N(\mu, \sigma^2/n)$

(b) $\overline{X}$ *and* $S^2$ *are independent*

(c) $(n-1)S^2/\sigma^2 \sim \chi^2(n-1)$

(d) *The random variable* $T = (\overline{X} - \mu)/(S/\sqrt{n}) \sim t(n-1)$