1.10.7 Bivariate Normal Distribution
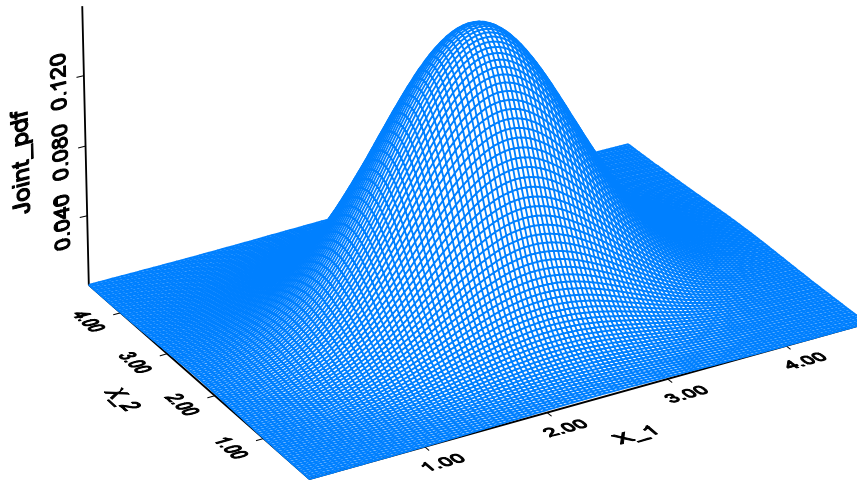

Figure 1.2: Bivariate Normal pdf

Here we use matrix notation. A bivariate rv is treated as a random vector

$$\boldsymbol{X} = \left(\begin{array}{c} X_1 \\ X_2 \end{array} \right).$$

The expectation of a bivariate random vector is written as

$$\boldsymbol{\mu} = \mathrm{E}\,\boldsymbol{X} = \mathrm{E} \left(\begin{array}{c} X_1 \\ X_2 \end{array} \right) = \left(\begin{array}{c} \mu_1 \\ \mu_2 \end{array} \right)$$

and its variance-covariance matrix is

$$\boldsymbol{V} = \left(\begin{array}{cc} \mathrm{var}(X_1) & \mathrm{cov}(X_1, X_2) \\ \mathrm{cov}(X_2, X_1) & \mathrm{var}(X_2) \end{array} \right) = \left(\begin{array}{cc} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{array} \right).$$

Then the joint pdf of a normal bi-variate rv $\boldsymbol{X}$ is given by

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{2\pi\sqrt{\det(\boldsymbol{V})}} \exp\left\{ -\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{V}^{-1}(\boldsymbol{x} - \boldsymbol{\mu}) \right\}, \qquad (1.18)$$

where $\boldsymbol{x} = (x_1, x_2)^{\mathrm{T}}$.

The determinant of $\boldsymbol{V}$ is

$$\det \boldsymbol{V} = \det \left(\begin{array}{cc} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{array} \right) = (1 - \rho^2)\sigma_1^2\sigma_2^2.$$

Hence, the inverse of V is

$$V^{-1} = \frac{1}{\det V} \begin{pmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{pmatrix} = \frac{1}{1-\rho^2} \begin{pmatrix} \sigma_1^{-2} & -\rho\sigma_1^{-1}\sigma_2^{-1} \\ -\rho\sigma_1^{-1}\sigma_2^{-1} & \sigma_2^{-2} \end{pmatrix}.$$

Then the exponent in formula (1.18) can be written as

$$-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^{\mathrm{T}} V^{-1}(\boldsymbol{x} - \boldsymbol{\mu}) =$$

$$= -\frac{1}{2(1-\rho^2)}(x_1 - \mu_1, x_2 - \mu_2) \begin{pmatrix} \sigma_1^{-2} & -\rho\sigma_1^{-1}\sigma_2^{-1} \\ -\rho\sigma_1^{-1}\sigma_2^{-1} & \sigma_2^{-2} \end{pmatrix} \begin{pmatrix} x_1 - \mu \\ x_2 - \mu \end{pmatrix}$$

$$= -\frac{1}{2(1-\rho^2)} \left(\frac{(x_1 - \mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1 - \mu_1)(x_2 - \mu_2)}{\sigma_1\sigma_2} + \frac{(x_2 - \mu_2)^2}{\sigma_2^2} \right).$$

So, the joint pdf of the two-dimensional normal rv $\boldsymbol{X}$ is

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{(1-\rho^2)}}$$

$$\times \exp\left\{ \frac{-1}{2(1-\rho^2)} \left(\frac{(x_1 - \mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1 - \mu_1)(x_2 - \mu_2)}{\sigma_1\sigma_2} + \frac{(x_2 - \mu_2)^2}{\sigma_2^2} \right) \right\}.$$

Note that when $\rho = 0$ it simplifies to

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \frac{1}{2\pi\sigma_1\sigma_2} \exp\left\{ -\frac{1}{2} \left(\frac{(x_1 - \mu_1)^2}{\sigma_1^2} + \frac{(x_2 - \mu_2)^2}{\sigma_2^2} \right) \right\},$$

which can be written as a product of the marginal distributions of X_1 and X_2. Hence, if $\boldsymbol{X} = (X_1, X_2)^{\mathrm{T}}$ has a bivariate normal distribution and $\rho = 0$ then the variables X_1 and X_2 are independent.

1.10.8 Bivariate Transformations

Theorem 1.17. *Let X and Y be jointly continuous random variables with joint pdf $f_{X,Y}(x, y)$ which has support on $\mathcal{S} \subseteq \mathbb{R}^2$. Consider random variables $U = g(X, Y)$ and $V = h(X, Y)$, where $g(\cdot, \cdot)$ and $h(\cdot, \cdot)$ form a one-to-one mapping from $\mathcal{S}$ to $\mathcal{D}$ with inverses $x = g^{-1}(u, v)$ and $y = h^{-1}(u, v)$ which have continuous partial derivatives. Then, the joint pdf of (U, V) is*

$$f_{U,V}(u, v) = f_{X,Y}\left(g^{-1}(u, v), h^{-1}(u, v) \right) |J|,$$

where, the Jacobian of the transformation J is

$$J = \det \begin{pmatrix} \frac{\partial g^{-1}(u,v)}{\partial u} & \frac{\partial g^{-1}(u,v)}{\partial v} \\ \frac{\partial h^{-1}(u,v)}{\partial u} & \frac{\partial h^{-1}(u,v)}{\partial v} \end{pmatrix}$$

for all $(u, v) \in \mathcal{D}$

□

Example 1.31. Let X, Y be independent rvs and $X \sim \text{Exp}(\lambda)$ and $Y \sim \text{Exp}(\lambda)$. Then, the joint pdf of (X, Y) is

$$f_{X,Y}(x, y) = \lambda e^{-\lambda x} \lambda e^{-\lambda y} = \lambda^2 e^{-\lambda(x+y)}$$

on support $\mathcal{S} = \{(x, y) : x > 0, y > 0\}$.

We will find the joint pdf for (U, V), where $U = g(X, Y) = X + Y$ and $V = h(X, Y) = X/Y$. This transformation and the support for (X, Y) give the support for (U, V). This is $\{(u, v) : u > 0, v > 0\}$.

The inverse functions are

$$x = g^{-1}(u, v) = \frac{uv}{1 + v} \quad \text{and } y = h^{-1}(u, v) = \frac{u}{1 + v}.$$

The Jacobian of the transformation is equal to

$$J = \det \begin{pmatrix} \frac{\partial g^{-1}(u,v)}{\partial u} & \frac{\partial g^{-1}(u,v)}{\partial v} \\ \frac{\partial h^{-1}(u,v)}{\partial u} & \frac{\partial h^{-1}(u,v)}{\partial v} \end{pmatrix} = \det \begin{pmatrix} \frac{v}{1+v} & \frac{u}{(1+v)^2} \\ \frac{1}{1+v} & -\frac{u}{(1+v)^2} \end{pmatrix} = \frac{-u}{(1+v)^2}.$$

Hence, by Theorem 1.17 we can write

$$f_{U,V}(u, v) = f_{X,Y}\left(g^{-1}(u, v), h^{-1}(u, v)\right)|J|$$

$$= \lambda^2 \exp\left\{-\lambda\left(\frac{uv}{1+v} + \frac{u}{1+v}\right)\right\} \times \frac{u}{(1+v)^2}$$

$$= \frac{\lambda^2 u e^{-\lambda u}}{(1+v)^2},$$

for $u, v > 0$.

□

These transformed variables are independent. In a simpler situation where $g(x)$ is a function of x only and $h(y)$ is function of y only, it is easy to see the following very useful result.

Theorem 1.18. *Let X and Y be independent rvs and let $g(x)$ be a function of x only and $h(y)$ be function of y only. Then the functions $U = g(X)$ and $V = h(Y)$ are independent.*

$\square$

Proof. (Continuous case) For any $u \in \mathbb{R}$ and $v \in \mathbb{R}$, define

$$A_u = \{x : g(x) \leq u\} \quad \text{and} \quad A_v = \{y : h(y) \leq v\}.$$

Then, we can obtain the joint cdf of (U, V) as follows

$$
\begin{aligned}
F_{U,V}(u, v) = P(U \leq u, V \leq v) &= P(X \in A_u, Y \in A_v) \\
&= P(X \in A_u)P(Y \in A_v) \quad \text{as } X \text{ and } Y \text{ are independent.}
\end{aligned}
$$

The mixed partial derivative with respect to u and v will give us the joint pdf for (U, V). That is,

$$f_{U,V}(u, v) = \frac{\partial^2}{\partial u \partial v} F_{U,V}(u, v) = \left(\frac{d}{du} P(X \in A_u) \right) \left(\frac{d}{dv} P(Y \in A_v) \right)$$

as the first factor depends on u only and the second factor on v only. Hence, the rvs $U = g(X)$ and $V = h(Y)$ are independent.

$\square$

Exercise 1.19. Let (X, Y) be a two-dimensional random variable with joint pdf

$$f_{X,Y}(x, y) = \begin{cases} 8xy, & \text{for } 0 \leq x < y \leq 1; \\ 0, & \text{otherwise.} \end{cases}$$

Let $U = X/Y$ and $V = Y$.

  (a) Are the variables X and Y independent? Explain.

  (b) Calculate the covariance of X and Y.

  (c) Obtain the joint pdf of (U, V).

  (d) Are the variables U and V independent? Explain.

  (e) What is the covariance of U and V?

Exercise 1.20. Let X and Y be independent random variables such that

$$X \sim \mathrm{Exp}(\lambda) \quad \text{and} \quad Y \sim \mathrm{Exp}(\lambda).$$

(a) Find the joint probability density function of (U, V), where

$$U = \frac{X}{X + Y} \quad \text{and} \quad V = X + Y.$$

(b) Are the variables U and V independent? Explain.

(c) Show that U is uniformly distributed on $(0, 1)$.

(d) What is the distribution of V?

1.11   Random Sample and Sampling Distributions

Example 1.32. In a study on relation of the level of cholesterol in blood and incidents of heart attack 28 heart attack patients had their cholesterol level measured two days and four days after the attack. Also, cholesterol levels were recorded for a control group of 30 people who did not have a heart attack. The data are available on the course web-page.

Various questions may be asked here. For example:

1. What is the population's mean cholesterol level on the second day after heart attack?

2. Is the difference in the mean cholesterol level on day 2 and on day 4 after the attack statistically significant?

3. Is high cholesterol level a significant risk factor for a heart attack?

Each numerical value in this example can be treated as a realization of a random variable. For example, value $x_1 = 270$ for patient one measured after two hours of the heart attack is a realization of a rv X_1 representing all possible values of cholesterol level of patient one. Here we have 28 patients, hence we have 28 random variables $X_1, \ldots, X_{28}$. These patients are only a small part (a sample) of all people (a population) having a heart attack (at this time and area of living). Here we come with a definition of a random sample.

Definition 1.21. *The random variables $X_1, \ldots, X_n$ are called a* random sample *of size n from a population if $X_1, \ldots, X_n$ are mutually independent, each having the same probability distribution.*

We say that such random variables are iid, that is *identically, independently distributed*.

The joint pdf (pmf in the discrete case) can be written as a product of the marginal pdfs, i.e.,

$$f_{\boldsymbol{X}}(x_1, \ldots, x_n) = \prod_{i=1}^{n} f_{X_i}(x_i),$$

where $\boldsymbol{X} = (X_1, \ldots, X_n)$ is the jointly continuous n-dimensional random variable.

Note: In the Example 1.32 it can be assumed in that these rvs are mutually independent (cholesterol level of one patient does not depend on the cholesterol level of another patient). If we can assume that the distribution of the cholesterol of each patient is the same, then the variables $X_1, \ldots, X_{28}$ constitute a random sample.

To make sensible inference from the observed values (a realization of a random sample) we often calculate some summaries of the data, such as the average or an estimate of the sample variance. In general, such summaries are functions of the random sample and we write

$$Y = T(X_1, \ldots, X_n).$$

Note that Y itself is then a random variable. The distribution of Y carries some information about the population and it allows us make inference regarding some parameters of the population. For example, about the expected cholesterol level and its variability in the population of people who suffer from heart attack.

The distributions of many functions $T(X_1, \ldots, X_n)$ can be derived from the distributions of the variables X_i. Such distributions are called *sampling distributions* and the function T is called a *statistic*.

Functions $\overline{X} = \frac{1}{n} \sum_{i-1}^{n} X_i$ and $S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2$ are the most common statistics used in data analysis.

1.11.1 χ_ν^2, t_ν and F_{ν_1, ν_2} Distributions

These three distributions can be derived from distributions of iid random variables. They are commonly used in statistical hypothesis testing and in interval estimation of unknown population parameters.

χ_ν^2 distribution

We have introduced the χ_ν^2 distribution as a special case of the gamma distribution, that is $\mathrm{Gamma}\left(\frac{\nu}{2}, \frac{1}{2}\right)$. We write

$$Y \sim \chi_\nu^2,$$

if the rv Y has the chi-squared distribution with ν degrees of freedom. ν is the parameter of the distribution function. The pdf of $Y \sim \chi_\nu^2$ is

$$f_Y(y) = \frac{y^{\frac{\nu}{2}-1}e^{-\frac{y}{2}}}{2^{\frac{\nu}{2}}\Gamma\left(\frac{\nu}{2}\right)} \quad \text{for } y > 0, \nu = 1, 2, \dots$$

and the mgf of Y is

$$M_Y(t) = \left(\frac{1}{1-2t}\right)^{\frac{\nu}{2}}, \quad t < \frac{1}{2}.$$

It is easy to show, using the derivatives of the mgf evaluated at $t = 0$, that

$$\mathrm{E}\,Y = \nu \quad \text{and} \quad \text{var}\,Y = 2\nu.$$

In the following example we will see that χ_ν^2 is the distribution of a function of iid standard normal rvs.

Example 1.33. In Example 1.17 we have seen that square of a standard normal rv has χ_1^2 distribution. Now, assume that $Z_1 \sim \mathcal{N}(0, 1)$ and $Z_2 \sim \mathcal{N}(0, 1)$ independently. What is the distribution of $Y = Z_1^2 + Z_2^2$? Denote by $Y_1 = Z_1^2$ and by $Y_2 = Z_2^2$.

To answer this question we can use the properties of the mgf as follows.

$$M_Y = \mathrm{E}\left(e^{tY}\right) = \mathrm{E}\left[e^{t(Y_1+Y_2)}\right] = \mathrm{E}\left[e^{tY_1}e^{tY_2}\right] = M_{Y_1}(t)M_{Y_2}(t)$$

as, by Theorem 1.18, Y_1 and Y_2 are independent. Also, $Y_1 \sim \chi_1^2$ and $Y_2 \sim \chi_1^2$, each with the mgf equal to

$$M_{Y_i}(t) = \frac{1}{(1-2t)^{1/2}}, \quad i = 1, 2.$$

Hence,

$$M_Y(t) = M_{Y_1}(t)M_{Y_2}(t) = \frac{1}{(1-2t)^{1/2}}\frac{1}{(1-2t)^{1/2}} = \frac{1}{(1-2t)}.$$

This is the mgf for χ_2^2. Hence, by the uniqueness of the mgf we can conclude that $Y = Z_1^2 + Z_2^2 \sim \chi_2^2$. $\qquad\square$

Note: This result can be easily extended to n independent standard normal rvs. That is, if $Z_i \sim \mathcal{N}(0, 1)$ for $i = 1, \dots, n$ independently, then

$$Y = \sum_{i=1}^{n} Z_i^2 \sim \chi_n^2.$$

From this result we can draw a useful conclusion.

Corollary 1.1. *A sum of independent random variables $T = \sum_{i=1}^{k} Y_i$, where each component has a chi-squared distribution, i.e., $Y_i \sim \chi^2_{\nu_i}$, is a random variable having a chi-squared distribution with $\nu = \sum_{i=1}^{k} \nu_i$ degrees of freedom.*

Note that T in the above corollary can be written as a sum of squared iid standard normal rvs, hence it must have a chi-squared distribution.

Example 1.34. Let $X_1, \ldots, X_n$ be iid random variables, such that

$$X_i \sim \mathcal{N}(\mu, \sigma^2), \quad i = 1, \ldots, n.$$

Then

$$Z_i = \frac{X_i - \mu}{\sigma} \sim \mathcal{N}(0, 1)$$

and so

$$\sum_{i=1}^{n} Z_i^2 = \sum_{i=1}^{n} \frac{(X_i - \mu)^2}{\sigma^2} \sim \chi_n^2.$$

This is a useful result, but it depends on two parameters, whose values are usually unknown (when we analyze data). Here we will see what happens when we replace μ with $\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$. We can write

$$\sum_{i=1}^{n} (X_i - \mu)^2 = \sum_{i=1}^{n} \left[(X_i - \overline{X}) + (\overline{X} - \mu) \right]^2$$

$$= \sum_{i=1}^{n} (X_i - \overline{X})^2 + \sum_{i=1}^{n} (\overline{X} - \mu)^2 + 2(\overline{X} - \mu) \underbrace{\sum_{i=1}^{n} (X_i - \overline{X})}_{=0}$$

$$= \sum_{i=1}^{n} (X_i - \overline{X})^2 + n(\overline{X} - \mu)^2.$$

Hence, dividing this by σ^2 we get

$$\sum_{i=1}^{n} \frac{(X_i - \mu)^2}{\sigma^2} = \frac{(n-1)S^2}{\sigma^2} + \frac{(\overline{X} - \mu)^2}{\frac{\sigma^2}{n}}, \qquad (1.19)$$

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

We know that (Intro to Stats)

$$\overline{X} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right), \quad \text{so} \quad \frac{\overline{X} - \mu}{\sqrt{\frac{\sigma^2}{n}}} \sim \mathcal{N}(0, 1).$$

Hence,

$$\left(\frac{\overline{X} - \mu}{\sqrt{\frac{\sigma^2}{n}}}\right)^2 \sim \chi_1^2.$$

Also

$$\sum_{i=1}^{n} \frac{(X_i - \mu)^2}{\sigma^2} \sim \chi_n^2.$$

Furthermore, it can be shown that $\overline{X}$ and S^2 are independent.

Now, equation (1.19) is a relation of the form $W = U + V$, where $W \sim \chi_n^2$ and $V \sim \chi_1^2$. Since here U and V are independent, we have

$$M_W(t) = M_U(t)M_V(t).$$

That is

$$M_U(t) = \frac{M_W(t)}{M_V(t)} = \frac{(1 - 2t)^{-n/2}}{(1 - 2t)^{-1/2}} = \frac{1}{(1 - 2t)^{(n-1)/2}}.$$

The last expression is the mgf of a random variable with a χ_{n-1}^2 distribution. That is

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2.$$

Equivalently,

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

$\square$

Student t_ν distribution

A derivation of the t-distribution was published in 1908 by William Sealy Gosset when he worked at the Guinness Brewery in Dublin. Due to proprietary issues, the paper was written under the pseudonym Student. The distribution is used in hypothesis testing and the test functions having a t distribution are often called t-tests.

We write

$$Y \sim t_\nu,$$

if the rv Y has the t distribution with ν degrees of freedom. ν is the parameter of the distribution function. The pdf is given by

$$f_Y(y) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{y^2}{\nu}\right)^{-\frac{\nu+1}{2}}, \quad y \in \mathbb{R}, \ \nu = 1, 2, 3, \ldots.$$

The following theorem is widely applied in statistics to build hypotheses tests.

Theorem 1.19. *Let Z and X be independent random variables such that*

$$Z \sim \mathcal{N}(0,1) \ \text{ and } \ X \sim \chi_\nu^2.$$

The random variable

$$Y = \frac{Z}{\sqrt{X/\nu}}$$

has Student t distribution with ν degrees of freedom.

Proof. Here we will apply Theorem 1.17. We will find the joint distribution of (Y, W), where $Y = \frac{Z}{\sqrt{X/\nu}}$ and $W = X$, and then we will find the marginal density function for Y, as required.

The densities of standard normal and chi-squared rvs, respectively, are

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \quad z \in \mathbb{R}$$

and

$$f_X(x) = \frac{x^{\frac{\nu}{2}-1} e^{-\frac{x}{2}}}{2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)}, \quad x \in \mathbb{R}_+$$

Z and X are independent, so the joint pdf of (Z, X) is equal to the product of the marginal pdfs $f_Z(z)$ and $f_X(x)$. That is,

$$f_{Z,X}(z, x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \frac{x^{\frac{\nu}{2}-1} e^{-\frac{x}{2}}}{2^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)}$$

Here we have the transformation

$$y = \frac{z}{\sqrt{x/\nu}} \ \text{ and } \ w = x,$$

which gives the inverses

$$z = y\sqrt{w/\nu} \ \text{ and } \ x = w.$$

Then, Jacobian of the transformation is

$$J = \det \begin{pmatrix} \sqrt{w/\nu} & \frac{y}{2\sqrt{\nu w}} \\ 0 & 1 \end{pmatrix} = \sqrt{\frac{w}{\nu}}.$$

Hence, the joint pdf for (Y, W) is

$$f_{Y,W}(y, w) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2 w}{2\nu}} \frac{w^{\frac{\nu}{2}-1} e^{-\frac{w}{2}}}{2^{\frac{\nu}{2}} \Gamma(\frac{\nu}{2})} \sqrt{\frac{w}{\nu}}$$

$$= \frac{w^{\frac{\nu+1}{2}-1} e^{-w\frac{1}{2}\left(1+\frac{y^2}{\nu}\right)}}{\sqrt{\nu\pi} 2^{\frac{\nu+1}{2}} \Gamma(\frac{\nu}{2})}.$$

The range of (Z, X) is $\{(z, x) : -\infty < z < \infty, 0 < x < \infty\}$ and so the range of (Y, W) is $\{(y, w) : -\infty < y < \infty, 0 < w < \infty\}$.

To obtain the marginal pdf for Y we need to integrate the joint pdf for (Y, W) over the range of W. This is an easy task when we notice similarities of this function with the pdf of $\mathrm{Gamma}(\alpha, \lambda)$ and use the fact that a pdf over the whole range integrates to 1.

The pdf of a gamma random variable V is

$$f_V(v) = \frac{v^{\alpha-1} \lambda^\alpha e^{-\lambda v}}{\Gamma(\alpha)}.$$

Denote

$$\alpha = \frac{\nu + 1}{2}, \quad \lambda = \frac{1}{2}\left(1 + \frac{y^2}{\nu}\right).$$

Then, multiplying and dividing the joint pdf for (Y, W) by λ^α and by $\Gamma(\alpha)$, we get

$$f_{Y,W}(y, w) = \frac{w^{\alpha-1} e^{-\lambda w}}{\sqrt{\nu\pi} 2^\alpha \Gamma(\frac{\nu}{2})} \times \frac{\lambda^\alpha \Gamma(\alpha)}{\lambda^\alpha \Gamma(\alpha)}$$

$$= \frac{w^{\alpha-1} \lambda^\alpha e^{-\lambda w}}{\Gamma(\alpha)} \times \frac{\Gamma(\alpha)}{\sqrt{\nu\pi} 2^\alpha \Gamma(\frac{\nu}{2}) \lambda^\alpha}.$$

Then, the marginal pdf for Y is

$$f_Y(y) = \int_0^\infty \frac{w^{\alpha-1} \lambda^\alpha e^{-\lambda w}}{\Gamma(\alpha)} \times \frac{\Gamma(\alpha)}{\sqrt{\nu\pi} 2^\alpha \Gamma(\frac{\nu}{2}) \lambda^\alpha} dw$$

$$= \frac{\Gamma(\alpha)}{\sqrt{\nu\pi} 2^\alpha \Gamma(\frac{\nu}{2}) \lambda^\alpha} \times \int_0^\infty \frac{w^{\alpha-1} \lambda^\alpha e^{-\lambda w}}{\Gamma(\alpha)} dw$$

as the second factor is treated as a constant when we integrate with respect to w. The first factor has a form of a gamma pdf, hence it integrates to 1. This gives the pdf for Y equal to

$$f_Y(y) = \frac{\Gamma(\alpha)}{\sqrt{\nu\pi}2^\alpha\Gamma(\frac{\nu}{2})\lambda^\alpha}$$

$$= \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}2^{\frac{\nu+1}{2}}\Gamma(\frac{\nu}{2})\left[\frac{1}{2}\left(1+\frac{y^2}{\nu}\right)\right]^{\frac{\nu+1}{2}}}$$

$$= \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})}\left(1+\frac{y^2}{\nu}\right)^{-\frac{\nu+1}{2}}$$

which is the pdf of a random variable having the Student t distribution with ν degrees of freedom. ☐

Example 1.35. Let X_i, $i = 1, \ldots, n$, be iid normal random variables with mean μ and variance σ^2. We often write this fact in the following way

$$X_i \underset{iid}{\sim} \mathcal{N}(\mu, \sigma^2), \quad i = 1, \ldots, n.$$

Then

$$\overline{X} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right), \quad \text{so } U = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim \mathcal{N}(0, 1).$$

Also, as shown in Example 1.34, we have

$$V = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}.$$

Furthermore, $\overline{X}$ and S^2 are independent, hence U and V are also independent (by Theorem 1.18). Hence, by Theorem 1.19, we have

$$T = \frac{U}{\sqrt{V/(n-1)}} \sim t_{n-1}.$$

That is

$$T = \frac{\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}}{\sqrt{\frac{(n-1)S^2}{\sigma^2}/(n-1)}} = \frac{\overline{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t_{n-1}$$

☐

Note: The function T in Example 1.35 is used for testing hypothesis about the population mean μ when the variance of the population is unknown. We will be using this function later on in this course.

F_{ν_1,ν_2} **Distribution**

Another sampling distribution very often used in statistics is the Fisher-Snedecor F distribution. We write

$$Y \sim F_{\nu_1,\nu_2},$$

if the rv Y has F distribution with ν_1 and ν_2 degrees of freedom. The degrees of freedom are the parameters of the distribution function. The pdf is quite complicated. It is given by

$$f_Y(y) = \frac{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)} \left(\frac{\nu_1}{\nu_2}\right) \frac{y^{\frac{\nu_1-2}{2}}}{\left(1+\frac{\nu_1}{\nu_2}y\right)^{\frac{\nu_1+\nu_2}{2}}}, \quad y \in \mathbb{R}_+.$$

This distribution is related to chi-squared distribution in the following way.

Theorem 1.20. *Let X and Y be independent random variables such that*

$$X \sim \chi^2_{\nu_1} \quad and \quad Y \sim \chi^2_{\nu_2}.$$

Then the random variable

$$F = \frac{X/\nu_1}{Y/\nu_2}$$

has Fisher's F distribution with ν_1 and ν_2 degrees of freedom.

Proof. This result can be shown in a similar way as the result of Theorem 1.19. Take the transformation $F = \frac{X/\nu_1}{Y/\nu_2}$ and $W = Y$ and use Theorem 1.17. $\square$

Example 1.36. Let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$ be two independent random samples such that

$$X_i \sim \mathcal{N}(\mu_1, \sigma_1^2) \quad and \quad Y_j \sim \mathcal{N}(\mu_2, \sigma_2^2), \quad i = 1, \ldots, n; \; j = 1, \ldots, m$$

and let S_1^2 and S_2^2 be the sample variances of the random samples $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$, respectively. Then (see Example 1.34),

$$\frac{(n-1)S_1^2}{\sigma_1^2} \sim \chi^2_{n-1}.$$

Similarly,

$$\frac{(m-1)S_2^2}{\sigma_2^2} \sim \chi^2_{m-1}.$$

Hence, by Theorem 1.20, the ratio

$$F = \frac{\frac{(n-1)S_1^2}{\sigma_1^2}/(n-1)}{\frac{(m-1)S_2^2}{\sigma_2^2}/(m-1)} = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F_{n-1,m-1}.$$

This statistic is used in testing hypotheses regarding variances of two independent normal populations.

□

Exercise 1.21. Let $Y \sim \text{Gamma}(\alpha, \lambda)$, that is, Y has the pdf given by

$$f_Y(y) = \begin{cases} \dfrac{y^{\alpha-1}\lambda^\alpha e^{-\lambda y}}{\Gamma(\alpha)}, & \text{for } y > 0; \\ 0, & \text{otherwise}, \end{cases}$$

for $\alpha > 0$ and $\lambda > 0$.

(a) Show that for any $a \in \mathbb{R}$ such that $a + \alpha > 0$ the expectation of Y^a is

$$\mathrm{E}(Y^a) = \frac{\Gamma(a + \alpha)}{\lambda^a \Gamma(\alpha)}.$$

(b) Obtain $\mathrm{E}(Y)$ and $\text{var}(Y)$. *Hint: Use result given in point (a) and the iterative property of gamma function, i.e., $\Gamma(z) = (z-1)\Gamma(z-1)$.*

(c) Let $X \sim \chi_\nu^2$. Use (1) to obtain $\mathrm{E}(X)$ and $\text{var}(X)$.

Exercise 1.22. Let $Y_1, \ldots, Y_5$ be a random sample of size 5 from a standard normal population, that is $Y_i \underset{iid}{\sim} \mathcal{N}(0, 1)$, $i = 1, \ldots, 5$, and denote by $\overline{Y}$ the sample mean, that is $\overline{Y} = \frac{1}{5}\sum_{i=1}^5 Y_i$. Let Y_6 be another independent standard normal random variable. What is the distribution of

(a) $W = \sum_{i=1}^5 Y_i^2$?

(b) $U = \sum_{i=1}^5 (Y_i - \overline{Y})^2$?

(c) $U + Y_6^2$?

(d) $\sqrt{5}Y_6/\sqrt{W}$?

Explain each of your answers.