ST 501: Fundamentals of Statistical Inference

Functions of random variables

Minh Tang

Department of Statistics, North Carolina State University.

Overview

Given a collection of random variables $X_1, X_2, \ldots, X_m$, the main focus of this lecture is to describe methods for deriving the distribution of $X = f(X_1, X_2, \ldots, X_m)$.

While the methods are applicable for general m and for any kind of random variables, in practice we will only consider the special cases of continuous random variables when $m = 1$ or $m = 2$. In the case when $m = 2$ we usually restrict ourselves to functions f that are either

$$f(X_1, X_2) = X_1 + X_2; \quad \text{or} \quad f(X_1, X_2) = X_1 - X_2$$
$$f(X_1, X_2) = X_1 X_2; \quad \text{or} \quad f(X_1, X_2) = X_1 / X_2.$$

Convolutions

We begin with the simplest transformation, namely addition or subtraction between two random variables.

Proposition
Let X and Y be discrete random variables with joint pmf p and marginal pmf p_X and p_Y. Let $Z = X + Y$. Then Z has pmf

$$p(z) = \sum_x p(x, z - x)$$

where the sum is over all x with $p_X(x) > 0$. If X and Y are independent then the above simplifies to

$$p(z) = \sum_x p_X(x) p_Y(z - x)$$

where the sum is once again over all x with $p_X(x) > 0$.

Proposition

Let X and Y be discrete random variables with joint pmf p and marginal pmf p_X and p_Y. Let $Z = X - Y$. Then X has pmf

$$p(z) = \sum_y p(z + y, y)$$

where the summation is over all y with $p_Y(y) > 0$. If X and Y are independent then the above simplifies to

$$p(z) = \sum_y p_X(z + y) p_y(y)$$

where the sum is once again over all y with $p_Y(y) > 0$.

Example: Trapezoidal distribution

Let X and Y be indepedent discrete uniform random variables on $\{1, \ldots, m\}$ and $\{1, \ldots, n\}$. Find the pmf for $Z = X + Y$.

Example: Trapezoidal distribution

Let X and Y be indepedent discrete uniform random variables on $\{1, \ldots, m\}$ and $\{1, \ldots, n\}$. Find the pmf for $Z = X + Y$.

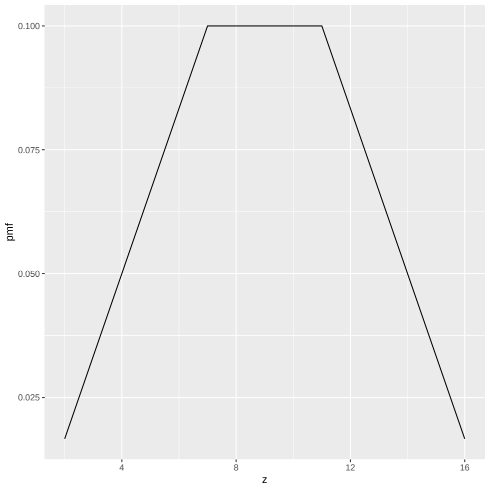
A. Assume first that $m \leq n$. Then $Z = X + Y$ has pmf

$$
\begin{aligned}
P(Z = z) &= \sum_x p_X(x) p_Y(z - x) = \sum_{x=1}^{m} \frac{1}{m} p_Y(z - x) \\
&= \sum_{x=1}^{m} \frac{1}{mn} I(1 \leq z - x \leq n) \\
&= \begin{cases} \frac{z-1}{mn} & 2 \leq z \leq m \\ \frac{1}{n} & m + 1 \leq z \leq n \\ \frac{m+n-z+1}{mn} & n + 1 \leq z \leq m + n \\ 0 & \text{otherwise} \end{cases}
\end{aligned}
$$

In the general case where possibly $m \geq n$, we replace m with $\min\{m, n\}$ and n with $\max\{m, n\}$, i.e.,

$$P(Z = z) = \begin{cases} \frac{z-1}{mn} & 2 \leq z \leq \min\{m, n\} \\ \frac{1}{\max\{m,n\}} & \min\{m, n\} + 1 \leq z \leq \max\{m, n\} \\ \frac{m+n-z+1}{mn} & \max\{m, n\} + 1 \leq z \leq m + n \\ 0 & \text{otherwise} \end{cases}$$

Example Let X be discrete uniform on $\{1, 2, \ldots, 6\}$ and Y be discrete uniform on $\{1, 2, \ldots, 10\}$ with X and Y independent. The pmf for $Z = X + Y$ is



Example: Binomial distribution

Let $X \sim \mathrm{Bin}(n,p)$ and $Y \sim \mathrm{Bin}(m,p)$ and X is independent of Y. Then $Z = X + Y$ has pmf

$$\begin{aligned}
P(Z=z) &= \sum_x p_X(x) p_Y(z-x) \\
&= \sum_{x=0}^{n} \binom{n}{x} p^x (1-p)^{n-x} \binom{m}{z-x} p^{z-x} (1-p)^{m-(z-x)} \\
&= p^z (1-p)^{n+m-z} \sum_{x=0}^{n} \binom{n}{x} \binom{m}{z-x} \\
&= \binom{n+m}{z} p^z (1-p)^{n+m-z}
\end{aligned}$$

where we have used Vandermonde identity.

Let $X \sim \mathrm{Bin}(n, p)$ and $Y \sim \mathrm{Bin}(m, p)$ and X is independent of Y. Then $Z = X - Y$ has pmf

$$\begin{aligned}
P(Z = z) &= \sum_{x} p_X(z + y) p_Y(y) \\
&= \sum_{y=0}^{m} \binom{n}{z+y} p^{z+y} (1-p)^{n-z-y} \binom{m}{y} p^y (1-p)^{m-y} \\
&= p^z (1-p)^{-z} \sum_{y=0}^{m} \binom{n}{z+y} \binom{m}{y} p^{2y} (1-p)^{n+m-2y}
\end{aligned}$$

which cannot be simplified further.

Convolutions for continuous rvs

Proposition

Let (X, Y) be a continuous bivariate random variable with joint pdf $f(x, y)$. Then $Z = X + Y$ has pdf

$$f_Z(z) = \int_{-\infty}^{\infty} f(x, z - x)\mathrm{d}x$$

When X and Y are independent then the above simplifies to

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z - x)\,\mathrm{d}x$$

where f_X and f_Y are the marginal pdf.

Proof First consider $P(Z \leq z)$, i.e.,

$$P(Z \leq z) = P(X + Y \leq z) = \int_{-\infty}^{\infty} \int_{-\infty}^{z-x} f(x, y) \, \mathrm{d}y \mathrm{d}x.$$

To get the pdf we differentiate $P(Z \leq z)$ with respect to z, i.e.,

$$\begin{aligned} f(z) &= \frac{\mathrm{d}}{\mathrm{d}z} \int_{-\infty}^{\infty} \int_{-\infty}^{z-x} f(x, y) \, \mathrm{d}y \mathrm{d}x \\ &= \int_{-\infty}^{\infty} \frac{\mathrm{d}}{\mathrm{d}z} \Big(\int_{-\infty}^{z-x} f(x, y) \, \mathrm{d}y \Big) \mathrm{d}x \\ &= \int_{-\infty}^{\infty} f(x, z - x) \, \mathrm{d}x \end{aligned}$$

as desired.

Proposition

Let (X, Y) be a continuous bivariate random variable with joint pdf $f(x, y)$. Then $Z = X - Y$ has pdf

$$f_Z(z) = \int_{-\infty}^{\infty} f(z+y, y)\mathrm{d}y$$

When X and Y are independent then the above simplifies to

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(z+y) f_Y(y)\,\mathrm{d}y$$

where f_X and f_Y are the marginal pdf.

Convolution of Normals

Let $X \sim \mathcal{N}(0,1)$, $Y \sim \mathcal{N}(0,c^2)$ with X and Y independent. Let $Z = X + Y$. Then

$$\begin{aligned} f(z) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\Big(-\frac{x^2}{2}\Big) \frac{1}{\sqrt{2\pi}c} \exp\Big(-\frac{(z-x)^2}{2c^2}\Big) \mathrm{d}x \\ &= \frac{1}{2\pi c} \int_{-\infty}^{\infty} \exp\Big(-\frac{x^2}{2} - \frac{z^2}{2c^2} + \frac{2xz}{2c^2} - \frac{x^2}{2c^2}\Big) \,\mathrm{d}x \\ &= \frac{1}{2\pi c} \exp\Big(-\frac{z^2}{2(1+c^2)}\Big) \int_{-\infty}^{\infty} \exp\Big(-\frac{a}{2}(x-bz)^2\Big) \mathrm{d}x \end{aligned}$$

where $a = \frac{1+c^2}{c^2}$ and $b = \frac{1}{1+c^2}$.

Do a change of variable, i.e.,

$$u = \sqrt{a}(x - bz); \quad \mathrm{d}x = \frac{1}{\sqrt{a}}\mathrm{d}u = \frac{c}{(1 + c^2)^{1/2}}\mathrm{d}u$$

we have

$$\begin{aligned} f(z) &= \exp\left(-\frac{z^2}{2(1 + c^2)}\right)\frac{1}{2\pi(1 + c^2)^{1/2}}\int_{-\infty}^{\infty} e^{-u^2/2}\,\mathrm{d}u \\ &= \frac{1}{\sqrt{2\pi}(1 + c^2)^{1/2}}\exp\left(-\frac{z^2}{2(1 + c^2)}\right) \end{aligned}$$

which is the pdf for a $\mathcal{N}(0, 1 + c^2)$ random variable.

Convolution of Gammas

Let $X \sim \Gamma(\alpha_1, \beta)$ and $Y \sim \Gamma(\alpha_2, \beta)$ be independent. Then $Z = X + Y$ has pdf (for $z \geq 0$)

Convolution of Gammas

Let $X \sim \Gamma(\alpha_1, \beta)$ and $Y \sim \Gamma(\alpha_2, \beta)$ be independent. Then $Z = X + Y$ has pdf (for $z \geq 0$)

$$
\begin{aligned}
f(z) &= \int_0^z \frac{1}{\Gamma(\alpha_1)\beta^{\alpha_1}} x^{\alpha_1 - 1} e^{-x/\beta} \frac{1}{\Gamma(\alpha_2)\beta^{\alpha_2}} (z - x)^{\alpha_2 - 1} e^{-(z-x)/\beta} \, \mathrm{d}x \\
&= \frac{1}{\Gamma(\alpha_1)\Gamma(\alpha_2)\beta^{\alpha_1+\alpha_2}} e^{-z/\beta} \int_0^z x^{\alpha_1 - 1}(z - x)^{\alpha_2 - 1} \, \mathrm{d}x \\
&= \frac{1}{\Gamma(\alpha_1)\Gamma(\alpha_2)\beta^{\alpha_1+\alpha_2}} e^{-z/\beta} z^{\alpha_1+\alpha_2-1} \int_0^1 t^{\alpha_1 - 1}(1 - t)^{\alpha_2 - 1} \, \mathrm{d}t \\
&= \frac{1}{\Gamma(\alpha_1)\Gamma(\alpha_2)\beta^{\alpha_1+\alpha_2}} e^{-z/\beta} z^{\alpha_1+\alpha_2-1} \frac{\Gamma(\alpha_1)\Gamma(\alpha_2)}{\Gamma(\alpha_1 + \alpha_2)} \\
&= \frac{1}{\Gamma(\alpha_1 + \alpha_2)\beta^{\alpha_1+\alpha_2}} z^{\alpha_1+\alpha_2-1} e^{-z/\beta}
\end{aligned}
$$

which is the pdf of a $\Gamma(\alpha_1 + \alpha_2, \beta)$ rv.

Coffee Break

Derive the pdf for $Z = X + Y$ when $X \sim \mathrm{Unif}(0, a)$ and $Y \sim \mathrm{Unif}(0, b)$ are independent.

Coffee Break

Derive the pdf for $Z = X + Y$ when $X \sim \mathrm{Unif}(0, a)$ and $Y \sim \mathrm{Unif}(0, b)$ are independent.

Note I hope that was a fun coffee break. This problem is quite a bit harder than it appears. The answer is, once again, a trapezoidal distribution, i.e.,

$$f_Z(z) = \begin{cases} \frac{z}{ab} & 0 \leq z \leq \min\{a, b\} \\ \frac{1}{\max\{a,b\}} & \min\{a, b\} \leq z \leq \max\{a, b\} \\ \frac{a+b-z}{ab} & \max\{a, b\} \leq z \leq a + b \\ 0 & \text{otherwise} \end{cases}$$

Example: Normal Poisson convolution

Let $Z = X + Y$ where $X \sim \mathrm{Pois}(\lambda)$ and $Y \sim \mathcal{N}(\mu, \sigma^2)$. Then Z is a continuous rv and

$$P(Z \le z) = \sum_x P(X = x) P(Y \le z - x) = \sum_{k=0}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!} \Phi((z - k - \mu)/\sigma)$$

The pdf of Z is then

$$\frac{\mathrm{d}}{\mathrm{d}z} P(Z \le z) = \sum_{k=0}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(z - \mu - k)^2}{2\sigma^2}\right)$$

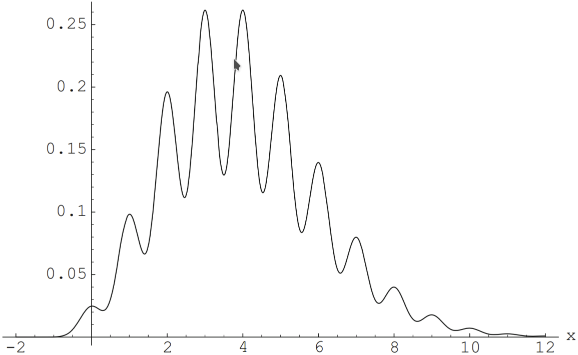


Fig. 4.1 Convolution of N(0, .09) and Poi (4)

Method of transformation

Recall the following transformation result that was used liberally in the earlier lectures.

Proposition
Let X be a continuous random variable with pdf f and let g be a function that is either monotone increasing or monotone decreasing. Then $Y = g(X)$ has pdf

$$f_Y(y) = \frac{f(g^{-1}(y))}{|g'(g^{-1}(y))|}$$

An interesting twist of this technique is its application to the case of bivariate random variables.

Let (X, Y) be a bivariate continuous r.v. with joint pdf $f_{XY}(x, y)$. Let $Z = g(X, Y)$ and suppose that, for all arbitrary but **fixed** x, $h(y) = g(x, y)$ is an increasing function of y.

Then conditional on $X = x$, the "conditional cdf" of Z is

$$
\begin{aligned}
P(Z \leq z \mid X = x) &= P(g(x, Y) \leq z \mid X = x) \\
&= P(h(Y) \leq z \mid X = x) \\
&= P(Y \leq h^{-1}(z) \mid X = x) \\
&= F_{Y|X}(h^{-1}(z) \mid x)
\end{aligned}
$$

The conditional pdf of Z given $X = x$ is then

$$
f_{Z|X}(z \mid x) = \frac{\mathrm{d}}{\mathrm{d}z} P(Z \leq z \mid X = x) = \frac{f_{Y|X}(h^{-1}(z) \mid x)}{|h'(h^{-1}(z))|}.
$$

Hence the joint pdf of (Z, X) is

$$f_{ZX}(z, x) = \frac{f_{Y|X}(h^{-1}(z) \mid x) f_X(x)}{|h'(h^{-1}(z))|} = \frac{f_{XY}(x, h^{-1}(z))}{|h'(h^{-1}(z))|}$$

The marginal pdf of Z is then

$$f_Z(z) = \int_{-\infty}^{\infty} f_{ZX}(z, x)\, \mathrm{d}x = \int_{-\infty}^{\infty} \frac{f_{XY}(x, h^{-1}(z))}{|h'(h^{-1}(z))|}\, \mathrm{d}x$$

The case when $g(x, y)$ is a decreasing function of y for each fixed x is derived similarly.

Example Let $Z = g(X, Y) = XY$.

For a fixed x, $y = h^{-1}(z) = z/x$ and $|h'(h^{-1}(z))| = |x|$. Hence

$$f_Z(z) = \int_{-\infty}^{\infty} \frac{f(x, z/x)}{|x|} \mathrm{d}x$$

Similarly, for $Z = X/Y$, by swapping the role of X and Y so that $x = h^{-1}(z) = yz$,

$$f_Z(z) = \int_{-\infty}^{\infty} |y| f(zy, y) \, \mathrm{d}y.$$

Example: Product and Quotient of Uniform rvs

Suppose $X \sim \mathrm{Unif}(0,1)$ and $Y \sim \mathrm{Unif}(0,1)$ are independent. Consider $U = XY$. Then $U \in [0,1]$ has pdf

$$f_U(u) = \int_0^1 \frac{f(x, u/x)}{|x|}\,\mathrm{d}x = \int_u^1 \frac{1}{x}\,\mathrm{d}x = -\ln u, \qquad 0 < u < 1$$

Meanwhile, for $V = X/Y \in [0,\infty)$,

$$f_V(v) = \int_0^1 |y| f(vy, y)\,\mathrm{d}y$$

Now if $v \le 1$ then $f(vy, y) > 0$ for all $y \in [0,1]$, i.e.,

$$f_V(v) = \int_0^1 |y| f(vy, y)\,\mathrm{d}y = \int_0^1 y\,\mathrm{d}y = \frac{1}{2}$$

Next if $v > 1$ then $f(vy, y) > 0$ provided $vy \le 1$, i.e.,

$$f_V(v) = \int_0^{1/v} |y| f(vy, y)\,\mathrm{d}y = \frac{1}{2v^2}$$

Example: Ratio of Normal rvs

Let X and Y be independent $\mathcal{N}(0, 1)$ and let $U = X/Y$. Then

$$
\begin{aligned}
f_U(u) &= \int_{-\infty}^{\infty} |y| f(vy, y)\,\mathrm{d}y \\
&= \int_{-\infty}^{\infty} |y| \frac{1}{2\pi} e^{-y^2/2} e^{-(vy)^2/2}\,\mathrm{d}y \\
&= \frac{1}{\pi} \int_{0}^{\infty} y e^{-y^2(1+v^2)/2} \mathrm{d}y \\
&= \frac{1}{\pi(1+v^2)}
\end{aligned}
$$

which is the pdf of a Cauchy rv.

24

Student's t distribution (optional)

We come to another widely abused distribution in statistics.

Let $X \sim \mathcal{N}(0,1)$ and $W \sim \chi^2_m$ be independent. Let $Z = \frac{X}{\sqrt{W/m}}$.

First $V = W/m$ has a chi-square distribution with shape parameters $\alpha = m/2$ and $\beta = 2/m$, i.e.,

$$f_V(v) = \frac{v^{m/2-1} e^{-mv/2}}{\Gamma(m/2)(2/m)^{m/2}} = \frac{m^{m/2} v^{m/2-1} e^{-mv/2}}{\Gamma(m/2) 2^{m/2}}$$

25

Now by the monotone transformation $Y = h(V) = \sqrt{V}$ has pdf

$$
\begin{aligned}
f_Y(y) &= \frac{f_V(h^{-1}(y))}{|h'(h^{-1}(y))|} \\
&= \frac{m^{m/2} y^{m-2} e^{-my^2/2}}{\Gamma(m/2) 2^{m/2}} \times (2y) \\
&= \frac{m^{m/2} y^{m-1} e^{-my^2/2}}{\Gamma(m/2) 2^{m/2-1}}
\end{aligned}
$$

Then $Z = \frac{X}{\sqrt{W/m}} = \frac{X}{Y}$ has pdf

$$
\begin{aligned}
f_Z(z) &= \int_0^\infty |y| f(zy, y)\,\mathrm{d}y \\
&= \frac{m^{m/2}}{\sqrt{2\pi}\Gamma(m/2)2^{m/2-1}} \int_0^\infty e^{-(zy)^2/2} y^m e^{-my^2/2}\,\mathrm{d}y \\
& \frac{m^{m/2}}{\sqrt{2\pi}\Gamma(m/2)2^{m/2-1}} \frac{\Gamma((m+1)/2)2^{(m-1)/2}}{(m+z^2)^{(m+1)/2}} \\
&= \frac{\Gamma((m+1)/2)}{\sqrt{m\pi}\Gamma(m/2)(1+z^2/m)^{(m+1)/2}}
\end{aligned}
$$

Note that $m = 1$ give the Cauchy pdf as expected.

27

As $m \to \infty$, the Student t "converges" to the standard normal.

Indeed, since $W = \chi_m^2$ is the sum of m independent χ_1^2, by Chebyshev's inequality, $W/m \longrightarrow \mathbb{E}[\chi_1^2] = 1$ and hence

$$V = \frac{X}{\sqrt{W/m}} \longrightarrow X \sim \mathcal{N}(0, 1)$$

Multivariable Transformations via Jacobian (Optional)

We now recall the following change of variable theorem of multivariable calculus. This theorem is of major importance in deriving the density of multivariate transformation.

The setup is as follows.

Let $X = (X_1, X_2, \ldots, X_n)$ have joint pdf $f(x_1, x_2, \ldots, x_n)$. Let $\mathcal{S}$ be the support of X, i.e., $P(X \in S) = 1$. Suppose $u_1, \ldots, u_n$ are n functions of $x_1, \ldots, x_n$, i.e., $u_i = g_i(x_1, x_2, \ldots, x_n)$ such that

- The map $x = (x_1, \ldots, x_n) \mapsto u = (u_1, , \ldots, u_n) \in \mathcal{T}$ is one-to-one for $x \in \mathcal{S}$
- The inverse functions $x_i = h_i(u_1, u_2, \ldots, u_n)$ are continuously differentiable on $\mathcal{T}$.

Assume also that the $n \times n$ Jacobian matrix

$$J(u) = \begin{bmatrix} \frac{\partial h_1}{\partial u_1} & \frac{\partial h_1}{\partial u_2} & \cdots & \frac{\partial h_1}{\partial u_n} \\ \frac{\partial h_2}{\partial u_1} & \frac{\partial h_2}{\partial u_2} & \cdots & \frac{\partial h_2}{\partial u_n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial h_n}{\partial u_1} & \frac{\partial h_n}{\partial u_2} & \cdots & \frac{\partial h_n}{\partial u_n} \end{bmatrix}$$

has determinant $|J(u)| \neq 0$ for all $(u_1, u_2, \ldots, u_n) \in \mathcal{T}$

Then the joint density of $U = (U_1, U_2, \ldots, U_n)$ is given by

$$f_U(u_1, u_2, \ldots, u_n) = f_X(h_1(u), h_2(u), \ldots, h_n(u)) \times |J(u)|$$

where $h_i(u) = h_i(u_1, u_2, \ldots, u_n)$.

We usually apply this result when the transformation g_i are sufficiently simple.

Example Let $X \sim \Gamma(\alpha_1, \beta)$ and $Y \sim \Gamma(\alpha_2, \beta)$ be independent. Find the distribution of $U = X/(X+Y)$

31

We introduce an auxiliary random variable $V = X + Y$ and note that $U \in [0, 1]$ and $V \in [0, \infty)$. The inverse transformation are given by

$$X = UV, \quad Y = (1 - U)V$$

The Jacobian matrix is then

$$J = \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix} = \begin{bmatrix} v & u \\ -v & (1 - u) \end{bmatrix}; \quad |J| = v$$

The joint pdf for (U, V) is then

$$
\begin{aligned}
f_{U,V}(u,v) &= f_{X,Y}(x,y) \times v \\
&= \frac{x^{\alpha_1-1}e^{-x/\beta}}{\Gamma(\alpha_1)\beta^{\alpha_1}} \frac{y^{\alpha_2-1}e^{-y/\beta}}{\Gamma(\alpha_2)\beta^{\alpha_2}} \times v \\
&= \frac{v(uv)^{\alpha_1-1}((1-u)v)^{\alpha_2-1}e^{-v/\beta}}{\Gamma(\alpha_1)\Gamma(\alpha_2)\beta^{\alpha_1+\alpha_2}} \\
&= \left(\frac{v^{\alpha_1+\alpha_2-1}e^{-v/\beta}}{\Gamma(\alpha_1+\alpha_2)\beta^{\alpha_1+\alpha_2}}\right) \times \left(\frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)}u^{\alpha_1-1}(1-u)^{\alpha_2-1}\right)
\end{aligned}
$$

The joint pdf for (U, V) factors into a product of the marginal pdf for $V \sim \Gamma(\alpha_1 + \alpha_2, \beta)$ and $U \sim \mathrm{Beta}(\alpha_1, \alpha_2)$.

U and V are thus independent

In the special case when $\alpha_1 = \alpha_2 = 1$, i.e., X and Y are independent exponential rv with mean β, then $U \sim \mathrm{Unif}(0, 1)$.

33

(Detour) Multivariate Normal Distribution

The Jacobian is essential in defining the multivariate normal one of the most important distribution in statistics.

Setup Let $X_1, \ldots, X_n$ be independent $\mathcal{N}(0,1)$ random variables. Let $\mathbf{A}$ be a $n \times n$ matrix. Let $U_1, \ldots, U_n$ be defined as

$$U_i = c_i + \sum_{j=1}^{n} a_{ij} X_j$$

where a_{ij} is the ijth entry of $\mathbf{A}$.

Note Each U_i is a linear function of the $\{X_j\}$. The map $X \mapsto U$ is one-to-one if and only if $\mathbf{A}$ is invertible.

Suppose that $\mathbf{A}$ is invertible and denote $\mathbf{B} = \mathbf{A}^{-1}$. The inverse transformation is then given by

$$X_i = \sum_j b_{ij}(U_j - c_j)$$

The joint pdf for U is then

$$f_U(u) = f_X(x_1, x_2, \ldots, x_n) \times |\mathbf{B}| = |\mathbf{A}|^{-1}(2\pi)^{-n/2} \prod_{i=1}^{n} e^{-x_i^2/2}$$

where $x = (x_1, x_2, \ldots, x_n) = \mathbf{A}^{-1}(u - c)$.

Example (bivariate normal) Let X_1 and X_2 be independent $\mathcal{N}(0,1)$. Define

$$U_1 = \sigma_1 X_1 + \mu_1$$

$$U_2 = \sigma_2(\rho X_1 + \sqrt{1-\rho^2} X_2) + \mu_2$$

Here $\rho \in [-1, 1]$. What is the joint pdf for $U = (U_1, U_2)$?

The inverse transformation for X_1 and X_2 are

$$X_1 = \frac{U_1 - \mu_1}{\sigma_1},$$

$$X_2 = \frac{U_2 - \mu_2}{\sigma_2\sqrt{1-\rho^2}} - \frac{\rho(U_1 - \sigma_1)}{\sigma_1\sqrt{1-\rho^2}}.$$

36

The matrix $\mathbf{B}$ is then

$$\mathbf{B} = \begin{bmatrix} \frac{1}{\sigma_1} & 0 \\ -\frac{\rho}{\sigma_1(1-\rho^2)^{1/2}} & \frac{1}{\sigma_2(1-\rho^2)^{1/2}} \end{bmatrix}; \qquad |\mathbf{B}| = \frac{1}{(1-\rho^2)^{1/2}\sigma_1\sigma_2}$$

The joint pdf for $U = (U_1, U_2)$ is then

$$\begin{aligned} f_U(u) &= \tfrac{1}{2\pi\sigma_1\sigma_2(1-\rho^2)^{1/2}} \exp(-x_1^2/2 - x_2^2/2) \\ &= \tfrac{1}{2\pi\sigma_1\sigma_2(1-\rho^2)^{1/2}} \exp\Big(\tfrac{-1}{1-\rho^2}\Big(\tfrac{(u_1-\mu_1)^2}{2\sigma_1^2} - \tfrac{\rho(u_1-\mu_1)(u_2-\mu_2)}{\sigma_1\sigma_2} + \tfrac{(u_2-\mu_2)^2}{2\sigma_2^2}\Big)\Big), \end{aligned}$$

quite a montrosity.

37

(Detour) Revisiting Student t

We next show one of the most fundamental result in "classical" applied statistics, i.e., that the sample mean for independent normal rvs is independent of the sample standard deviation.

Proposition
Let $X_1, X_2, \ldots, X_n$ be iid $\mathcal{N}(\mu, \sigma^2)$ rvs. Define

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2.$$

Then $\bar{X}$ and S^2 are independent.

38

Proof Let $U_1 = \bar{X} - \mu$ and define, for $i = 2, \ldots, n$

$$U_i = X_i - \bar{X}$$

The inverse transformations (a linear transformation) are then

$$X_i - \mu = U_i + U_1, \quad i = 2, 3, \ldots, n; \qquad X_1 - \mu = U_1 - \sum_{i=2}^{n} U_i.$$

The joint pdf for $U = (U_1, U_2, \ldots, U_n)$ is then

$$\begin{aligned}
f_U(u) &= \frac{|J|}{(2\pi\sigma^2)^{n/2}} \prod_{i=1}^{n} \exp\Big(-\frac{(x_i - \mu)}{2\sigma^2}\Big) \\
&= \frac{|J|}{(2\pi\sigma^2)^{n/2}} \exp\Big(-\tfrac{(u_1 - \sum_{i=2}^{n} u_i)^2}{2\sigma^2}\Big) \prod_{i=2}^{n} \exp\Big(-\tfrac{(u_i + u_1)^2}{2\sigma^2}\Big) \\
&= \frac{|J|}{(2\pi\sigma^2)^{n/2}} \exp\Big(-\frac{n u_1^2}{2\sigma^2}\Big) \exp\Big(-\frac{(\sum_{i=2}^{n} u_i)^2 + \sum_{i=2}^{n} u_i^2}{2\sigma^2}\Big)
\end{aligned}$$

Important $f_U(u)$ is the product of a function that depends only on u_1 and another function that depends on $(u_2, \ldots, u_n)$.

In other words, u_1 is independent of $v = (u_2, u_3, \ldots, u_n)$.

We note that the sample variance can be written as

$$\begin{aligned}(n-1)S^2 = \sum_{i=1}^{n}(X_i - \bar{X})^2 &= (X_1 - \bar{X})^2 + \sum_{i=2}^{n}(X_i - \bar{X})^2 \\ &= (U_2 + U_3 + \cdots + U_n)^2 + \sum_{i=2}^{n} U_i^2\end{aligned}$$

and so S^2 is a function of the $U_2, U_3, \ldots, U_n$ only.

We conclude that $\bar{X} = U_1 + \mu$ is independent of S^2 as desired.

It is now a simple exercise to show that

Theorem
Let $X_1, \ldots, X_n$ be iid $\mathcal{N}(\mu, \sigma^2)$. Then $\bar{X}$ and S^2 are independent and $(n-1)S^2/\sigma^2 \sim \chi^2_{n-1}$.

Proof We start with the simple observation that

$$
\begin{aligned}
\underbrace{\sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma} \right)^2}_{W} &= \sum_{i=1}^{n} \left(\frac{X_i - \bar{X} + \bar{X} - \mu}{\sigma} \right)^2 \\
&= \sum_{i=1}^{n} \left(\frac{X_i - \bar{X}}{\sigma} \right)^2 + n \left(\frac{\bar{X} - \mu}{\sigma} \right)^2 \\
&= \underbrace{\frac{(n-1)S^2}{\sigma^2}}_{Y} + \underbrace{\frac{n(\bar{X} - \mu)^2}{\sigma^2}}_{Z}
\end{aligned}
$$

We first note that $W \sim \chi^2_n$ and $Z \sim \chi^2_1$. Furthermore, since S^2 is independent of $\bar{X}$, we have $m_W(t) = m_Y(t) \times m_Z(t)$, i.e.,

$$m_Y(t) = \frac{m_W(t)}{m_Z(t)} = \frac{(1-2t)^{-n/2}}{(1-2t)^{-1/2}} = (1-2t)^{-(n-1)/2}$$

By the uniqueness of mgf, $Y = (n-1)S^2/\sigma^2 \sim \chi^2_{n-1}$ as desired.

We have thus shown that the statistic

$$\frac{\sqrt{n}(\bar{X}-\mu)}{S} = \frac{\sqrt{n}(\bar{X}-\mu)/\sigma}{\sqrt{((n-1)S^2/\sigma^2)/(n-1)}} = \frac{\mathcal{N}(0,1)}{\sqrt{\chi^2_{n-1}/(n-1)}}$$

is a Student t with $n-1$ degrees of freedom.

42