A Few Special Distributions and Their Properties

Econ 690

Purdue University

Special Distributions and Their Associated Properties

The Uniform Distribution

A continuous random variable, Y, has a *Uniform distribution* over the interval $[a, b]$, denoted $Y \sim U(a, b)$, if its p.d.f. is given by:

$$f_U(y|a, b) = \begin{cases} \frac{1}{b-a} & \text{if } \ a \leq y \leq b \\ 0 & \text{otherwise,} \end{cases}$$

where $-\infty < a < b < \infty$.

If $Y \sim U(a, b)$ then $E(Y) = \frac{a+b}{2}$ and $var(Y) = \frac{(b-a)^2}{12}$.

The Gamma Distribution

A continuous random variable Y has a *Gamma distribution* with mean $\mu > 0$ and degrees of freedom $\nu > 0$, denoted $Y \sim \gamma(\mu, \nu)$, if its p.d.f. is:

$$f_\gamma(y|\mu,\nu) \equiv \begin{cases} c_\gamma^{-1} y^{\frac{\nu-2}{2}} \exp\left(-\frac{y\nu}{2\mu}\right) & \text{if } 0 < y < \infty \\ 0 & \text{otherwise,} \end{cases}$$

where the integrating constant is given by $c_\gamma = \left(\frac{2\mu}{\nu}\right)^{\frac{\nu}{2}} \Gamma\left(\frac{\nu}{2}\right)$.
It is also common to parameterize the Gamma in terms of $\alpha = \frac{\nu}{2}$ and $\beta = \frac{2\mu}{\nu}$, in which case we denote the distribution as $Y \sim G(\alpha, \beta)$. The associated density function is denoted by $f_G(y|\alpha,\beta)$ where

$$f_G(y|\alpha,\beta) \equiv \begin{cases} c_G^{-1} y^{\alpha-1} \exp(-y/\beta) & \text{if } 0 < y < \infty \\ 0 & \text{otherwise,} \end{cases}$$

and $c_G = \beta^\alpha \Gamma(\alpha)$.

Mean and Variance of the Gamma Distribution

If $Y \sim G(\alpha, \beta)$ then $E(Y) = \alpha\beta$ and $\text{Var}(Y) = \alpha\beta^2$. If $Y \sim \gamma(\mu, \nu)$, then $E(Y) = \mu$ and $\text{Var}(Y) = 2\mu^2/\nu$.

Notes: Distributions related to the Gamma include the *Chi-square distribution* which is a Gamma distribution with $\nu = \mu$. It is denoted by $Y \sim \chi^2(\nu)$. The *Exponential distribution* is a Gamma distribution with $\nu = 2$.

The Inverse Gamma Distribution

We denote the *inverted Gamma* density as $Y \sim IG(\alpha, \beta)$. Though different parameterizations exist (particularly for how β enters the density), we utilize the following form here:

$$Y \sim IG(\alpha, \beta) \Rightarrow p(y) = [\Gamma(\alpha)\beta^{\alpha}]^{-1} y^{-(\alpha+1)} \exp(-1/[y\beta]), \quad y > 0.$$

The mean of this inverse Gamma is $E(Y) = [\beta(\alpha-1)]^{-1}$, for $\alpha > 1$, and

the variance is $\text{Var}(Y) = [\beta^2(\alpha-1)^2(\alpha-2)]^{-1}$ for $\alpha > 2$.

The Multivariate Normal Distribution

A continuous $k-$dimensional random vector, $Y = (Y_1, .., Y_k)'$, has a *Normal distribution* with mean μ (a $k-$vector) and variance Σ (a $k \times k$ positive definite matrix), denoted $Y \sim N(\mu, \Sigma)$, if its p.d.f. is given by:

$$\phi(y|\mu, \Sigma) = \phi(y; \mu, \Sigma) = \frac{1}{(2\pi)^{\frac{k}{2}}} |\Sigma|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(y-\mu)' \Sigma^{-1} (y-\mu)\right].$$

The cumulative distribution function of the multivariate Normal, evaluated at the point y^*, is denoted by $\Phi(y^*|\mu, \Sigma)$ or, if $\mu = 0, \Sigma = I$, by $\Phi(y^*)$
Note: The special case where $k = 1$, $\mu = 0$ and $\Sigma = 1$ is referred to as the *Standard Normal* distribution.

Marginals and Conditionals of Multivariate Normal

Suppose the $k-$vector $Y \sim N(\mu, \Sigma)$ is partitioned as:

$$Y = \begin{pmatrix} Y_{(1)} \\ Y_{(2)} \end{pmatrix}$$

where $Y_{(i)}$ is a k_i-vector for $i = 1, 2$ with $k_1 + k_2 = k$ and μ and Σ have been partitioned conformably as:

$$\mu = \begin{pmatrix} \mu_{(1)} \\ \mu_{(2)} \end{pmatrix}$$

and

$$\Sigma = \begin{pmatrix} \Sigma_{(11)} & \Sigma_{(12)} \\ \Sigma'_{(12)} & \Sigma_{(22)} \end{pmatrix}.$$

Marginals and Conditionals, Continued

Then the following results hold:

- The marginal distribution of $Y_{(i)}$ is $N\left(\mu_{(i)}, \Sigma_{(ii)}\right)$ for $i = 1, 2$.
- The conditional distribution of $Y_{(1)}$ given $Y_{(2)} = y_{(2)}$ is $N\left(\mu_{(1|2)}, \Sigma_{(1|2)}\right)$ where

$$\mu_{(1|2)} = \mu_{(1)} + \Sigma_{(12)}\Sigma_{(22)}^{-1}\left(y_{(2)} - \mu_{(2)}\right)$$

and

$$\Sigma_{(1|2)} = \Sigma_{(11)} - \Sigma_{(12)}\Sigma_{(22)}^{-1}\Sigma_{(12)}'.$$

The Multivariate Student-t

A continuous $k-$dimensional random vector, $Y = (Y_1, .., Y_k)'$, has a *t distribution* with mean μ (a $k-$vector), scale matrix Σ (a $k \times k$ positive definite matrix) and ν (a positive scalar referred to as a *degrees of freedom* parameter), denoted $Y \sim t(\mu, \Sigma, \nu)$, if its p.d.f. is given by:

$$f_t(y|\mu, \Sigma, \nu) = \frac{1}{c_t}|\Sigma|^{-\frac{1}{2}} \left[\nu + (y-\mu)' \Sigma^{-1} (y-\mu)\right]^{-\frac{\nu+k}{2}},$$

where

$$c_t = \frac{\pi^{\frac{k}{2}} \Gamma\left(\frac{\nu}{2}\right)}{\nu^{\frac{\nu}{2}} \Gamma\left(\frac{\nu+k}{2}\right)}.$$

Notes: The special case where $k = 1$, $\mu = 0$ and $\Sigma = 1$ is referred to as the *Student-t* distribution with ν degrees of freedom. Tables providing percentiles of the Student-t are available in most econometrics and statistics textbooks. The case where $\nu = 1$ is referred to as the *Cauchy distribution*.

Mean and Variance of Student-t

If $Y \sim t(\mu, \Sigma, \nu)$ then $E(Y) = \mu$ if $\nu > 1$ and $var(Y) = \frac{\nu}{\nu - 2}\Sigma$ if $\nu > 2$.

Notes: The mean and variance only exist if $\nu > 1$ and $\nu > 2$, respectively. This implies, for instance, that the mean of the Cauchy does not exist even though it is a valid p.d.f. and, hence, its median and other quantiles exist. Σ is not exactly the same as the variance matrix and, hence, is given another name: the *scale matrix*.

Marginals and Conditionals of Student-t

Suppose the $k-$vector $Y \sim t(\mu, \Sigma, \nu)$ is partitioned as in our description of the multivariate Normal, as are μ and Σ. Then the following results hold:

- The marginal distribution of $Y_{(i)}$ is $t\left(\mu_{(i)}, \Sigma_{(ii)}, \nu\right)$ for $i = 1, 2$.
- The conditional distribution of $Y_{(1)}$ given $Y_{(2)} = y_{(2)}$ is $t\left(\mu_{(1|2)}, \Sigma_{(1|2)}, \nu + k_1\right)$ where

$$\mu_{(1|2)} = \mu_{(1)} + \Sigma_{(12)}\Sigma_{(22)}^{-1}\left(y_{(2)} - \mu_{(2)}\right),$$

$$\Sigma_{(1|2)} = h_{(1|2)}\left[\Sigma_{(11)} - \Sigma_{(12)}\Sigma_{(22)}^{-1}\Sigma_{(12)}'\right]$$

and

$$h_{(1|2)} = \frac{1}{\nu + k_2}\left[\nu + \left(y_{(2)} - \mu_{(2)}\right)'\Sigma_{(22)}^{-1}\left(y_{(2)} - \mu_{(2)}\right)\right].$$

The Wishart Distribution

Let H be an $N \times N$ positive definite (symmetric) random matrix, A be a fixed (non-random) $N \times N$ positive definite matrix and $\nu > 0$ a scalar degrees of freedom parameter. Then H has a *Wishart distribution*, denoted $H \sim W(A, \nu)$, if its p.d.f. is given by:

$$f_W(H|A,\nu) = \frac{1}{c_W} |H|^{\frac{\nu-N-1}{2}} |A|^{-\frac{\nu}{2}} \exp\left[-\frac{1}{2} tr\left(A^{-1}H\right)\right],$$

where

$$c_W = 2^{\frac{\nu N}{2}} \pi^{\frac{N(N-1)}{4}} \prod_{i=1}^{N} \Gamma\left(\frac{\nu+1-i}{2}\right).$$

Note: If $N = 1$, then the Wishart reduces to a *Gamma distribution* *(i.e.* $f_W(H|A,\nu) = f_G(H|\nu A,\nu)$ *if* $N = 1$*)*.

Some Moments of the Wishart Distribution

If $H \sim W(A, \nu)$ then

$$E(H_{ij}) = \nu A_{ij},$$

$$var(H_{ij}) = \nu \left(A_{ij}^2 + A_{ii} A_{jj} \right), \quad i, j = 1, .., N$$

and

$$cov(H_{ij}, H_{km}) = \nu \left(A_{ik} A_{jm} + A_{im} A_{jk} \right), \quad i, j, k, m = 1, .., N,$$

where subscripts i, j, k, m refer to elements of matrices.

The Binomial Distribution

A discrete random variable, Y, has a *Binomial distribution* with parameters T and p, denoted $Y \sim B(T, p)$, if its probability function is given by:

$$f_B(y|T,p) = \begin{cases} \frac{T!}{(T-y)!y!} p^y (1-p)^{T-y} & \text{if } \ y = 0, 1, .., T \\ 0 & \text{otherwise,} \end{cases}$$

where $0 \leq p \leq 1$ and T is a positive integer. The *Bernoulli* distribution is a special case of the Binomial when $T = 1$.
If $Y \sim B(T, p)$ then

$$E(Y) = Tp, \quad var(Y) = Tp(1-p).$$

Note: This distribution is used in cases where an experiment, the outcome of which is either "success" or "failure", is repeated independently T times. The probability of success in an experiment is p. The distribution of the random variable Y, which counts the number of successes, is $B(T, p)$.

The Poisson Distribution

A discrete random variable, Y, has a *Poisson distribution* with parameter λ, denoted $Y \sim Po(\lambda)$, if its probability function is given by:

$$f_{Po}(y|\lambda) = \begin{cases} \frac{\lambda^y \exp(-\lambda)}{y!} & \text{if } \; y = 0, 1, 2, \ldots \\ 0 & \text{otherwise,} \end{cases}$$

where λ a positive real number.

If $Y \sim Po(\lambda)$ then $E(Y) = \lambda$ and $var(Y) = \lambda$.

The Multinomial Distribution

A discrete $N-$dimensional random vector, $Y = (Y_1, .., Y_N)'$, has a *Multinomial distribution* with parameters T and p, denoted $Y \sim M(T, p)$, if its probability density function is given by:

$$f_M(y|T,p) = \begin{cases} \frac{T!}{y_1!..y_N!} p_1^{y_1}..p_N^{y_N} & \text{if } \; y_i = 0, 1, .., T \text{ and } \sum_{i=1}^N y_i = T \\ 0 & \text{otherwise,} \end{cases}$$

where $p = (p_1, .., p_N)'$, $0 \leq p_i \leq 1$ for $i = 1, .., N$, $\sum_{i=1}^N p_i = 1$ and T is a positive integer.

The Dirichlet and Beta Distributions

Let $Y = (Y_1, .., Y_N)'$ be a vector of continuous random variables with the property that $Y_1 + .. + Y_N = 1$. Then Y has a *Dirichlet distribution*, denoted $Y \sim D(\alpha)$, if its p.d.f. is given by:

$$f_D(Y|\alpha) = \left[\frac{\Gamma(a)}{\prod_{i=1}^N \Gamma(\alpha_i)}\right] \prod_{i=1}^N y_i^{\alpha_i - 1},$$

where $\alpha = (\alpha_1, .., \alpha_N)'$, $\alpha_i > 0$ for $i = 1, .., N$ and $a = \sum_{i=1}^N \alpha_i$. The *Beta distribution*, denoted by $Y \sim \beta(\alpha_1, \alpha_2)$, is the Dirichlet distribution for the case $N = 2$. Its p.d.f. is denoted by $f_B(Y|\alpha_1, \alpha_2)$.

Note: In the case $N = 2$, the restriction $Y_1 + Y_2 = 1$ can be used to remove one of the random variables. Thus, the Beta distribution is a univariate distribution.

Moments of the Dirichlet Distribution

Suppose $Y \sim D(\alpha)$ where α and a are as given as in the previous page. Then for $i, j = 1, .., N$,

- $E(Y_i) = \frac{\alpha_i}{a}$,
- $var(Y_i) = \frac{\alpha_i(a-\alpha_i)}{a^2(a+1)}$ and
- $cov(Y_i, Y_j) = -\frac{\alpha_i \alpha_j}{a^2(a+1)}$.

The Pareto Distribution

A continuous random variable Y has a *Pareto distribution* if its p.d.f. is given by:

$$f_{Pa}(y|\gamma,\lambda) = \begin{cases} \frac{\lambda\gamma^{\lambda}}{y^{\lambda+1}} & \text{if } \; y \geq \gamma \\ 0 & \text{otherwise} \end{cases} .$$