18.440 Probability and Random Variables
Spring 2009

Gamma and beta probabilities

This handout is based on section 1.5 of a book manuscript, *Handbook and "Tables" of Classic Probabilities*, by Robert J. Holt, R. M. Dudley, David Yang Gao, and Lewis Pakula. The numbering from that section is preserved, but some revisions have been made.

1.5 Gamma and beta functions and probabilities. The *gamma function* is defined for any $a > 0$ by

$$(1.5.1) \qquad \Gamma(a) := \int_0^\infty x^{a-1} e^{-x} dx.$$

The integral is finite if (and only if) $a > 0$, because $\int_0^1 x^{a-1} dx = 1/a < \infty$, and $x^{a-1} < e^{x/2}$ for x large enough.

Integration by parts shows that $\Gamma(a+1) = a\Gamma(a)$ for any $a > 0$. We have $\Gamma(1) = 1$. It follows by induction that $\Gamma(k+1) = k!$ for any nonnegative integer k.

For any $a > 0$ the function defined by

$$(1.5.2) \qquad \gamma_a(x) := x^{a-1} e^{-x} / \Gamma(a)$$

for $x > 0$, and 0 for $x \leq 0$, is a probability density. The corresponding distribution is called a *gamma distribution with parameter* a.

If the random variable X has a gamma distribution with parameter a then $EX = a$ since $EX = \Gamma(a+1)/\Gamma(a)$. Likewise $EX^2 = \Gamma(a+2)/\Gamma(a) = (a+1)a$ so $\mathrm{Var}(X) = a$ and $\sigma_X = a^{1/2}$.

Recall that for any random variable X with density f and any $c > 0$, cX has a density $c^{-1} f(x/c)$. Applying that to $c = 1/\lambda$ for any $\lambda > 0$, if X has density γ_a then X/λ has the density $\gamma_{a,\lambda}$ defined by

$$\gamma_{a,\lambda}(x) = \lambda^a x^{x-1} e^{-\lambda x} / \Gamma(a)$$

for $0 < x < +\infty$ and 0 otherwise. A random variable Y with this density evidently has $EY = a/\lambda$ and $\mathrm{Var}(Y) = a/\lambda^2$.

The *Beta function* is defined for any $a > 0$ and $b > 0$ by

$$(1.5.3) \qquad B(a,b) := \int_0^1 x^{a-1}(1-x)^{b-1} dx.$$

Clearly, $0 < B(a,b) < \infty$ for any $a > 0$ and $b > 0$. Letting $y := 1 - x$ shows that $B(b,a) \equiv B(a,b)$. Let $\beta_{a,b}(x) := x^{a-1}(1-x)^{b-1}/B(a,b)$ for $0 < x < 1$ and 0 for $x \leq 0$ or $x \geq 1$. Then $\beta_{a,b}$ is a probability density. The probability distribution with this density is called a *beta* distribution with parameters a, b. Its distribution function is then defined as

$$(1.5.4) \qquad I_x(a,b) := \int_0^x \beta_{a,b}(t) dt, \quad 0 \leq x \leq 1.$$

The following fact relates gamma distributions with different parameters with each other and relates gamma and beta functions.

1.5.5 Theorem. For any $a > 0$ and $b > 0$,
(a) $B(a, b) \equiv B(b, a) \equiv \Gamma(a)\Gamma(b)/\Gamma(a + b)$.
(b) If X and Y are independent random variables having gamma distributions with parameters a and b respectively, then $U := X + Y$ has a gamma distribution with parameter $a + b$.

Proof. First consider (b). U has a density u given by a convolution of those of X and Y, namely, for any $x > 0$,

$$\begin{aligned} u(x) &= \int_0^x \gamma_a(x-y)\gamma_b(y)dy \\ &= \int_0^x (x-y)^{a-1}e^{-(x-y)}y^{b-1}e^{-y}dy/(\Gamma(a)\Gamma(b)) \\ &= e^{-x}\int_0^x (x-y)^{a-1}y^{b-1}dy/(\Gamma(a)\Gamma(b)). \end{aligned}$$

The substitution $y = tx, \quad 0 \le t \le 1$ gives

$$= e^{-x}x^{a+b-1}B(b, a)/(\Gamma(a)\Gamma(b)).$$

Since u must be a probability density, it must be the gamma density with parameter $a+b$, and the normalizing constants must agree, so both (a) and (b) follow. □

Iterating Theorem 1.5.5, it follows that if X_i are independent identically distributed variables, each having the standard exponential distribution with density e^{-x} for $x \ge 0$ and 0 for $x < 0$, so that the X_i have gamma distributions with parameter 1, then for each $n = 1, 2, ..., S_n = X_1 + ... + X_n$ has a γ_n density. If each X_i has a $\gamma_{a,\lambda}$ density then S_n has a $\gamma_{na,\lambda}$ density.

It is now easy to find the means and variances of beta distributions. If X has a beta distribution with parameters a, b, in other words has distribution function (1.5.4), then $EX = B(a+1, b)/B(a, b)$. Similarly $EX^2 = B(a+2, b)/B(a, b) = a(a+1)/[(a+b)(a+b+1)]$. Thus

$$EX = a/(a + b), \quad \mathrm{Var}(X) = \frac{ab}{(a + b)^2(a + b + 1)}. \tag{1.5.6}$$

Note that $1 - X$ has a beta distribution with parameters b, a. Thus $E(1 - X) = b/(a + b)$ which equals $1 - a/(a + b)$ as it should. Also, $1 - X$ has the same variance as X, and so the expression for $\mathrm{Var}(X)$ is preserved by interchanging a and b as it should be.

Let $0 < \lambda < \infty$ and let Y be a Poisson random variable with parameter λ. Then some notations are

$$P(k, \lambda) = \Pr(Y \le k) = e^{-\lambda}\sum_{j=0}^{k} \lambda^j/j!,$$

$$Q(k, \lambda) \;=\; \Pr(Y \geq k) \;=\; e^{-\lambda} \sum_{j=k}^{\infty} \lambda^j / j!.$$

There are identities relating the Poisson and gamma distributions:

1.5.7 Theorem. For any positive integer k, if X has a γ_k density, we have for any $x \geq 0$,

$$Q(k, x) = P(X \leq x) \tag{1.5.8}$$

and

$$P(k-1, x) = P(X > x). \tag{1.5.9}$$

For $0 < \lambda < \infty$, if Y has a $\gamma_{k,\lambda}$ density and $0 < t < \infty$, then

$$P(Y \leq t) = Q(k, \lambda t) \tag{1.5.10}$$

and

$$P(Y > t) P(k-1, \lambda t). \tag{1.5.11}$$

Proof. Equation (1.5.9) follows by differentiating with respect to x and noting that the derivative of $P(k-1, x)$ gives a (finite) telescoping sum. Equation follows by taking complements.

Then letting $Y = X/\lambda$, Y has the given density, and (1.5.11) follows from (1.5.9), and (1.5.10) follows by taking complements or from (1.5.8). $\square$

A similar identity relates beta and binomial probabilities. Let $0 < p < 1$, $q = 1 - p$, let X be a binomial (n, p) random variable and

$$B(k, n, p) \;=\; \Pr(X \leq k) \;=\; \sum_{j=0}^{k} b(j, n, p),$$

$$E(k, n, p) \;=\; \Pr(X \geq k) \;=\; \sum_{j=k}^{n} b(j, n, p).$$

1.5.12 Theorem. If $0 < p < 1$, and $0 \leq k \leq n$ are integers, then

$$E(k, n, p) \;=\; I_p(k, n-k+1), \quad \text{if} \quad k \geq 1; \quad B(k, n, p) \;=\; I_{1-p}(n-k, k+1), \quad \text{if} \quad k < n.$$

Proof. The first equality again follows from differentiating a finite sum with respect to p which gives a telescoping sum. The second then follows from $B(k, n, p) \equiv E(n-k, n, 1-p)$. $\square$

Quotients of independent variables with the densities just given have distributions that can be expressed in terms of beta probabilities:

1.5.13 Theorem. Let X and Y be independent gamma variables with parameters (a, λ) and (b, μ) respectively. Then for $0 < t < \infty$, $P(Y/X \leq t) = P(V \geq \lambda/(\lambda + \mu t))$ where V has a beta distribution with parameters a, b.

Proof. We have

$$P(Y \leq tX) = \int_0^\infty \lambda^a x^{a-1} e^{-\lambda x} \Gamma(a)^{-1} \int_0^{tx} \mu^b y^{b-1} e^{-\mu y} \Gamma(b)^{-1} dy dx.$$

For each fixed x, make the substitution $y = sx$ in the inner integral. Then

$$P(Y \leq tX) = \int_0^\infty \lambda^a x^{a-1} e^{-\lambda x} \Gamma(a)^{-1} \int_0^{t} \mu^b (sx)^{b-1} e^{-\mu s x} \Gamma(b)^{-1} x ds dx.$$

The integral is absolutely convergent, so the integrals with respect to s and x can be interchanged. We have

$$\int_0^\infty x^{a+b-1} e^{-(\lambda+\mu s)x} dx = \Gamma(a+b)/(\lambda + \mu s)^{a+b}.$$

Thus

$$P(Y \leq tX) = \lambda^a \mu^b \int_0^t s^{b-1} (\lambda + \mu s)^{-a-b} ds / B(a, b).$$

Make the substitution $v := \lambda/(\lambda + \mu s)$. Then

$$P(Y \leq tX) = \int_{\lambda/(\lambda+\mu t)}^1 v^{a-1}(1-v)^{b-1} dv / B(a, b).$$

$\square$

.