Exact calculation of inequality probabilities

John D. Cook
Department of Biostatistics
P. O. Box 301402, Unit 1409
The University of Texas, M. D. Anderson Cancer Center
Houston, Texas 77230-1402, USA
cook@mdanderson.org

November 16, 2009

Abstract

This note surveys results for computing the inequality probability

$$P(X > Y)$$

in closed form where X and Y are independent continuous random variables. Distribution families discussed include

- normal,
- Cauchy,
- gamma,
- inverse gamma,
- Lévy,
- folded normal, and
- beta.

Mixture distributions are also discussed.

1 Normal and Cauchy random variables

The probability

$$P(X > Y)$$

can be evaluated simply when X and Y are either both normal or both Cauchy random variables. These probabilities are derived in [1].

If X is normal with mean μ_X and variance σ_X^2, and Y is normal with mean μ_Y and variance σ_Y^2, then

$$P(X > Y) = \Phi\left(\frac{\mu_X - \mu_Y}{(\sigma_X^2 + \sigma_Y^2)^{1/2}}\right)$$

where $\Phi(x)$ is the CDF of a standard normal random variable.

If X is Cauchy with location μ_X and scale σ_X, and Y is Cauchy with location μ_Y and scale σ_Y, then

$$P(X > Y) = \frac{1}{2} + \frac{1}{\pi}\tan^{-1}\left(\frac{\mu_X - \mu_Y}{\sigma_X + \sigma_Y}\right).$$

2 Gamma, inverse gamma, and Lévy random variables

If X has a gamma(α_X, β_X) distribution and Y has a gamma(α_Y, β_Y) distribution then

$$P(X > Y) = P\left(B < \frac{\beta_X}{\beta_X + \beta_Y}\right) = I_{\beta_X/(\beta_X+\beta_Y)}(\alpha_Y, \alpha_X)$$

where $B \sim \text{beta}(\alpha_Y, \alpha_X)$ and $I_x(a, b)$ is the regularized incomplete beta function. This result first appeared in [1].

If X has an inverse gamma distribution, $1/X$ has a gamma distribution. Therefore inverse gamma inequality probabilities can be reduced to gamma inequality probabilities since

$$P(X > Y) = P(1/Y > 1/X).$$

A Lévy distribution with scale c is the same as an inverse gamma distribution with shape $1/2$ and scale $c/2$, and so inequalities with Lévy distributions are a special case of inequalities with inverse gamma distributions.

(There are multiple conventions for parameterizing gamma and inverse gamma distributions. The results here assume the density of a gamma(α, β) random variable is proportional to $x^{\alpha-1}\exp(-x/\beta)$. Also, the density of an inverse gamma(α, β) random variable is proportional to $x^{-\alpha-1}\exp(-\beta/x)$.)

3 Folded normal random variables

A folded normal random variable is the absolute value of a normal random variable. Therefore folded normal random inequalities reduce to computing $P(|X| > |Y|)$ where X and Y are normal.

Suppose X is normal(μ_X, σ_X^2) and Y is normal(μ_Y, σ_Y^2). If $\sigma_X^2 = \sigma_Y^2$ then $P(|X| > |Y|)$ can be computed as follows. Let $U = X + Y$ and $V = X - Y$. Then

$$\begin{aligned} U &\sim N(\mu_U, \sigma_U^2) = N(\mu_X + \mu_Y, \sigma_X + \sigma_Y) \\ V &\sim N(\mu_V, \sigma_V^2) = N(\mu_X - \mu_Y, \sigma_X + \sigma_Y). \end{aligned}$$

When $\sigma_X^2 = \sigma_Y^2$, we have

$$\Pr(|X| > |Y|) = \Phi\left(-\frac{\mu_U}{\sigma_U}\right)\Phi\left(-\frac{\mu_V}{\sigma_V}\right) + \Phi\left(\frac{\mu_U}{\sigma_U}\right)\Phi\left(\frac{\mu_V}{\sigma_V}\right). \tag{1}$$

See [4] for a derivation of the above result as well as a method for evaluating $P(|X| > |Y|)$ when the variances of X and Y are not equal.

4 Beta random variables

In general $P(X > Y)$ cannot be evaluated in closed form if X and Y have beta distributions. However, many special cases do have closed form expressions. Let $X \sim \text{beta}(a, b)$ and $Y \sim \text{beta}(c, d)$. Define

$$g(a, b, c, d) = P(X > Y).$$

The function $g(a, b, c, d)$ can be evaluated in closed form if

- one of the four parameters a, b, c, or d is an integer,
- the fractional parts of the parameters sum to 1, or
- $a + b$ and $c + d$ are positive integers.

The symmetries

$$g(a, b, c, d) = g(d, c, b, a) = g(d, b, c, a) = 1 - g(c, d, a, b)$$

are developed in [1]. For the 24 permutations of the 4 arguments, there are at most six different values of g. These are $g(a,b,c,d)$, $g(a,b,d,c)$, $g(a,c,d,b)$ and their complementary probabilities.

If we define

$$h(a,b,c,d) = \frac{B(a+c,b+d)}{B(a,b)B(c,d)} \tag{2}$$

$$= \frac{\Gamma(a+c)\Gamma(b+d)\Gamma(a+b)\Gamma(c+d)}{\Gamma(a)\Gamma(b)\Gamma(c)\Gamma(d)\Gamma(a+b+c+d)}. \tag{3}$$

then [1] shows that the following recurrence relations hold:

$$\begin{aligned}
g(a+1,b,c,d) &= g(a,b,c,d) + h(a,b,c,d)/a \\
g(a,b+1,c,d) &= g(a,b,c,d) - h(a,b,c,d)/b \\
g(a,b,c+1,d) &= g(a,b,c,d) - h(a,b,c,d)/c \\
g(a,b,c,d+1) &= g(a,b,c,d) + h(a,b,c,d)/d.
\end{aligned}$$

The identity

$$g(a,b,c,1) = \frac{\Gamma(a+b)\,\Gamma(a+c)}{\Gamma(a+b+c)\,\Gamma(a)}$$

is derived in [2]. The symmetries and recurrence relationships above can be used to bootstrap this result into a closed form evaluation of g when any one of its arguments is an integer.

If $a+b+c+d=1$, then [2] shows that

$$g(a,b,c,d) = \frac{\sin(\pi a)\sin(\pi d)}{\sin(\pi(a+b))\sin(\pi(b+d))}.$$

Symmetry and recurrence relationships can extend this result to compute g for any arguments whose fractional parts sum to 1.

Finally, if $a+b=c+d=1$ and $a+c \neq 1$ then

$$g(a,b,c,d) = \frac{\Gamma(a+c)\Gamma(b+d)\sin(\pi a1)\sin(\pi b_2)}{\pi^2}\frac{c(\psi(c)-\psi(1-a))}{a+c-1}$$

where ψ is the digamma function. As before, this result can be extended by using symmetry and recurrence relationships.

For more information regarding beta inequalities, see [2] and [3].

5 Mixture distributions

If $P(X > Y)$ can be computed in closed form for distributions X and Y from a particular family, then it can also be computed in closed form for mixtures of distributions from that family.

Suppose

$$f_X = \sum_{i=1}^{n} \lambda_i f_{X_i}$$

and

$$f_Y = \sum_{j=1}^{m} \theta_j f_{Y_j}$$

where $\lambda_i \geq 0$, $\theta_i \geq 0$, and

$$\sum_{i=1}^{n} \lambda_i = \sum_{j=1}^{m} \theta_j = 1.$$

Then

$$P(X > Y) = \sum_{i=1}^{n} \sum_{j=1}^{m} \lambda_i \theta_j P(X_i > Y_j).$$

References

[1] John D. Cook. Numerical computation of stochastic inequality probabilities (2003). UT MD Anderson Cancer Center Department of Biostatistics Working Paper Series. Working Paper 46.

[2] John D. Cook. Exact calculation of beta inequalities (2005). Technical Report UTMDABTR-005-05.

[3] John D. Cook and Saralees Nadarajah. Stochastic Inequality Probabilities for Adaptively Randomized Clinical Trials. Biometrical Journal. 48 (2006) pp 256-365.

[4] John D. Cook. Inequality Probabilities for Folded Normal Random Variables (2009). UT MD Anderson Cancer Center Department of Biostatistics Working Paper Series. Working Paper 52.