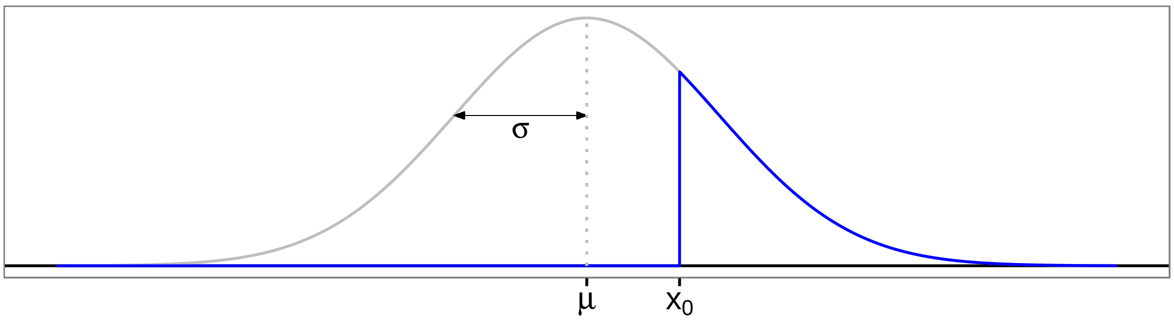
Truncated Normal Proposal

Bret Larget

4/15/2017



Truncated normal density proposal.

We wish to find a density function proportional to the right tail of a normal distribution in order to match a likelihood function which is defined only on positive values and the log of the likelihood is very well approximated by a parabola. Relating to the graph above, the gray curve is a Normal(μ, σ^2) density $g(x)$, but we seek a distribution F with density $f(x)$ proportional to $g(x)1_{\{x>x_0\}}$. We will set $x_0 = 0$. We wish to find values of μ and σ in order to match specified a specified mean and variance of the distribution F. We also require an expression for the inverse cdf of F for random number generation.

Density

The density f has the following expression.

$$f(x) = \frac{1}{(1 - \Phi(z_0))\sigma\sqrt{2\pi}} \, \mathrm{e}^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \qquad x > x_0$$

where Φ is the cdf of the standard normal curve and $z_0 = (x_0 - \mu)/\sigma$.

Cumulative distribution function

The cdf of F has the following expression.

$$F(x) = \frac{\Phi\left(\frac{x-\mu}{\sigma}\right) - \Phi(z_0)}{1 - \Phi(z_0)}$$

Here is a derivation.

$$
\begin{aligned}
F(x) &= 1 - \int_x^\infty \frac{1}{(1-\Phi(z_0))\sigma\sqrt{2\pi}}\, \mathrm{e}^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2}\, \mathrm{d}t \\
&= 1 - \frac{1}{1-\Phi(z_0)} \int_{\frac{x-\mu}{\sigma}}^\infty \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{z^2}{2}}\, \mathrm{d}z \\
&= 1 - \frac{1-\Phi(\frac{x-\mu}{\sigma})}{1-\Phi(z_0)} \\
&= \frac{\Phi\left(\frac{x-\mu}{\sigma}\right) - \Phi(z_0)}{1-\Phi(z_0)}, \qquad x > x_0
\end{aligned}
$$

Random Generation

Using the inverse cdf method, we can generate a random variable with distribution F by the following equation where $U \sim \text{Uniform}(0,1)$.

$$X = \mu + \sigma\Phi^{-1}(1 - (1-U)(1-\Phi(z_0)))$$

The derivation is straightforward.

Moments

It remains to find values of μ and σ so that the distribution F has a given mean and standard deviation. Define $\mu_X = \mathsf{E}(X)$ and $\sigma_X = \sqrt{\mathsf{Var}(X)}$ when $X \sim F$. An expression for μ_X is as follows.

$$\mu_X = \mu + \frac{\sigma}{\sqrt{2\pi}(1-\Phi(z_0))}\, \mathrm{e}^{-\frac{z_0^2}{2}} = \mu + \sigma h(z_0) \qquad \text{where } h(z_0) = \frac{\phi(z_0)}{1-\Phi(z_0)}$$

In survival analysis, $h(t)$ is called the *hazard rate* of a probability density.

An expression for $\mathsf{E}(X^2)$ is

$$\mathsf{E}(X^2) = \mu^2 + \sigma^2 + \frac{(2\mu + z_0)\sigma \mathrm{e}^{-z_0^2/2}}{\sqrt{2\pi}(1-\Phi(z_0))}$$

The variance is (should be a simpler expression!)

$$\sigma_X^2 = \sigma^2(1 - h(z_0)^2) + \sigma z_0 h(z_0)$$

Solving for μ and σ

For a given μ_X and σ_X^2, we want to find values of μ and σ to achieve these moments when $x_0 = 0$.

Here, $z_0 = (x_0 - \mu)/\sigma = -\mu/\sigma$ so that $\mu = -z_0\sigma$. Reparameterizing in terms of z_0 and σ may be easier. The equations we need to solve for z_0 and σ are:

$$
\begin{aligned}
\mu_X &= \sigma(h(z_0) - z_0) \\
\sigma_X^2 &= \sigma^2(1 - (h(z_0))^2) + \sigma z_0 h(z_0)
\end{aligned}
$$