Computational Statistics
1. exercise session, 29.1.2010

1. (More conjugate analysis.) We observe the number of deaths y_i during one year due to a certain disease in a number of cities $i = 1, \ldots, n$ with known populations $x_i \times 100000$ people. We are interested in the (disease-specific) death rate θ expressed as cases per year and per 100000 people.

We model the data as the observed values of random variables Y_i, where

$$[Y_i \mid \Theta = \theta] \sim \mathrm{Poi}(x_i\,\theta), \qquad i = 1, \ldots, n,$$

independently. The prior is $\mathrm{Gam}(a, b)$ with known parameters $a > 0$ and $b > 0$.

a) Give a formula for the likelihood.

b) Find the posterior distribution.

2. Pareto distribution with parameters $k, a > 0$ has the cdf

$$F(x) = \begin{cases} 1 - \left(\dfrac{k}{x}\right)^a, & x \geq k \\ 0, & x < k. \end{cases}$$

How does one simulate this distribution using the inverse transform?

3. Let X be a random variable with discrete distrubution and the pmf

$$f(x) = \begin{cases} 0.3, & \text{when } x = 0, \\ 0.2, & \text{when } x = 1, \\ 0.5, & \text{when } x = 2. \end{cases}$$

Calculate the distribution function and draw it. (It should be a right continuous step function.) Further, calculate the quantile function (using the formula for the generalized inverse function) and draw it. (It should also be a step function. Is it right continuous or left continuos?)

Check that the inverse transform gives the obvious simulation algorithm: we break the unit interval into three subintervals with lengths 0.3, 0.2 and 0.5, which we label 0, 1, 2, respectively. We simulate $U \sim \mathrm{Uni}(0, 1)$ and then return the label of that subinterval, into which U falls.

4. Suppose that the random variable $X > 0$ has a continuous distribution. Give an expression for f_Y in terms of f_X, when

a) $Y = X^{1/r}/b$, where $b, r > 0$.

b) $Y = \ln X$.

5. Let U be a continuous random variable with $0 < U < 1$ and define X as the logit of U, $X = \text{logit}(U)$, where

$$\text{logit}(p) = \ln \frac{p}{1-p}, \qquad 0 < p < 1.$$

Determine an expression for the density of X in terms of the density of U.

Remarks: (1) If p is the probability of an event happening, then the **odds** of that event happening (as opposed it not happening) are $p/(1-p)$, so $\text{logit}(p)$ is the same as log-odds. (2) If $U \sim \text{Uni}(0,1)$, then X has the (standard) logistic distribution.