Lab 2

January 22, 2013

HW1

	Length n=30	Length n=5	Coverage n=30	Coverage n=5
Frequentist	0.277	0.510	0.945	0.664
Uniform	0.272	0.566	0.960	0.942
Beta(8,2)	0.238	0.374	0.978	0.999

Table: Simulation resutls.

Model for count data

- $y_i \in \{0, 1, 2, \ldots, \}$: # of emails subject i received yesterday ($i = 1, \ldots, n$).
- Likelihood [Poisson(θ)]:

$$\begin{aligned} L\,(y\,|\,\theta) &= \prod_{i=1}^{n} \frac{\theta^{y_i}}{y_i!} e^{-\theta}, \\ &= \frac{1}{\prod_i y_i!} \theta^{\sum_{i=1}^{n} y_i} e^{-n\theta} \end{aligned}$$

- Prior [Gamma(α,β)]:

$$\pi(\theta) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\beta\theta}.$$

- Posterior:

$$\begin{aligned}
\pi(\theta \,|\, y) &\propto L\,(y\,|\,\theta)\pi(\theta), \\
&\propto \theta^{\sum_{i=1}^{n} y_i} e^{-n\theta} \theta^{\alpha-1} e^{-\beta\theta}, \\
&= \theta^{\sum_{i=1}^{n} y_i + \alpha - 1} e^{-(n+\beta)\theta}, \\
&\sim \text{Gamma}\left(\sum_{i=1}^{n} y_i + \alpha, n + \beta\right).
\end{aligned}$$

Truncated random variables

- Guess $\theta \in [a, b]$ with probability one.
- $f(\theta)$: density for θ.
- $f_{[a,b]}(\theta)$: truncated density on $[a, b]$,

$$f_{[a,b]}(\theta) = \frac{f(\theta)1_{(a\leq\theta\leq b)}}{\int_a^b f(z)dz} = \frac{f(\theta)1_{(a\leq\theta\leq b)}}{F(b) - F(a)},$$

where $F(x) = P[\theta \leq x]$ is cdf and

$$1_{(a\leq\theta\leq b)} = \begin{cases} 1, & a \leq \theta \leq b, \\ 0, & \text{otherwise.} \end{cases}$$

Example: Truncated Beta

- Consider Beta(α,β) truncated to $[0.4, 0.6]$.
- The truncated beta density at point p can be written in R,

$$
\begin{aligned}
\text{dbeta}(p, \alpha, \beta) * \text{as.numeric}(p > 0.4 \ \&\ p < 0.6) \\
/(\text{pbeta}(0.6, \alpha, \beta) - \text{pbeta}(0.4, \alpha, \beta))
\end{aligned}
$$

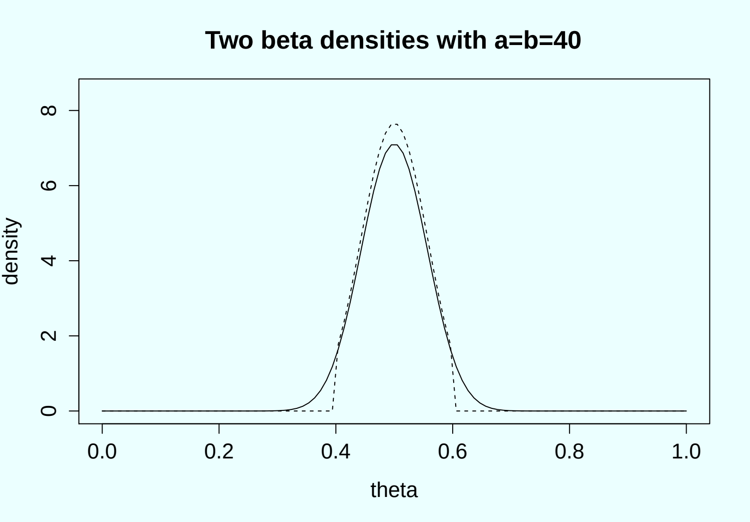

Figure: Truncated and untruncated beta(40,40) densities.

Truncated random variables

- $$\int f_{[a,b]}(\theta)d\theta = \int \frac{f(\theta)1_{(a\le\theta\le b)}}{\int_a^b f(z)dz}d\theta = \frac{\int_a^b f(\theta)d\theta}{\int_a^b f(z)dz} = 1$$
- CDF of truncated random variables

$$\begin{aligned} F_{[a,b]}(z) &= \int_{-\infty}^{z} f_{[a,b]}(\theta)d\theta = \int_{-\infty}^{z} \frac{f(\theta)1_{(a\le\theta\le b)}}{F(b)-F(a)}d\theta, \\ &= \frac{F(z)-F(a)}{F(b)-F(a)}. \end{aligned}$$

Truncated random variables

- Computing a p-quantile point. We want to know z such that

$$F_{[a,b]}(z) = p,$$
$$\frac{F(z) - F(a)}{F(b) - F(a)} = p.$$

Then,

$$z = F^{-1}\left[\{F(b) - F(a)\}p + F(a)\right].$$

- If $f(\theta)$ is beta, F^{-1}: qbeta and F: pbeta in R code.

R: writing own functions

- $f(x) = x + 2$:

  ```
  func1 = function(x){
  y = x + 2
  return(y)
  }
  func1(3)
  5
  ```

- $f(x, y) = 2x - y$:

  ```
  func2 = function(x, y){
  z = 2 * x - y
  return(z)
  }
  func2(2,2)
  2
  ```

- Let x = (x[1], ..., x[n]) and y = (y[1], ..., y[n]).
- sapply:
  sapply(x, func1) returns
  func1(x[1]), ..., func1(x[n])
  sapply(1:3, func1)
  3, 4, 5
- mapply:
  mapply(func2, x, y) returns
  func2(x[1],y[1]), ..., func2(x[n],y[n])
  mapply(func2, 1:3, 2:4)
  4, 7, 10
  mapply(func2, 1:3, 2)
  func2(1,2), func2(2,2), func2(3,2)
  4, 6, 8