

## 5.61

$$(a) \quad f_Y(y) = \frac{1}{B(a, b)} y^{a-1} (1-y)^{b-1} \quad 0 \leq y \leq 1$$

$$f_V(y) = \frac{1}{B([a], [b])} y^{[a]-1} (1-y)^{[b]-1}$$

$$\frac{f_Y(y)}{f_V(y)} = \frac{B(a, b)}{B([a], [b])} y^{a-[a]} (1-y)^{b-[b]} \quad 0 \leq y \leq 1$$

$$\because a - [a] \in [0, 1), \quad b - [b] \in [0, 1)$$

$$\therefore y^{a-[a]}, (1-y)^{b-[b]} \in [0, 1]$$

即  $M = \sup \frac{f_Y(y)}{f_V(y)}$  有界。

(b) 这小问题目有问题，仅当 V 的参数  $b'$  大于  $b$ ，或者  $a'$  等于  $a$  时，M 才为有限值。如果将 V 的参数改为  $([a], b+1)$  则

$$f_Y(y) = \frac{1}{\Gamma(a)b^a} y^{a-1} e^{-\frac{y}{b}}$$

$$f_V(y) = \frac{1}{\Gamma([a])b^{[a]}} y^{[a]-1} e^{-\frac{y}{b+1}}$$

$$\frac{f_Y(y)}{f_V(y)} = \frac{\Gamma([a])(b+1)^{[a]}}{\Gamma(a)b^a} y^{a-[a]} e^{-y(\frac{1}{b} - \frac{1}{b+1})}$$

(c) 若将 V 的参数换成  $[a] + 1$ ，(a)，(b) 中  $y$  对应的指数  $a - [a] - 1$  变为负，于是当  $y \rightarrow 0$  时， $\frac{f_Y(y)}{f_V(y)} \rightarrow \infty$ ，即 M 会变为无穷大。

(d) 如果仅仅是为了使  $EN$  最小，显然当 V, Y 的分布一样时，M 最小为 1。但是考虑到问题的实际背景，即为了通过简单的容易产生的随机数来产生复杂的随机数。而由前面的例题可以知道，可以用均匀分布随机变量较容易的产生出两个参数都为整数的 Beta 分布以及  $\alpha$  参数为整数的 Gamma

分布。因此在寻求使  $EN$  最小的参数时，应当将范围限制在上述允许的条件内。

对 Beta 分布，设  $V$  的参数为  $(a', b')$ ，则

$$\frac{f_Y(y)}{f_V(y)} = \frac{B(a, b)}{B(a', b')} y^{a-a'} (1-y)^{b-b'}$$

记  $\alpha = a - a'$ ， $\beta = b - b'$ ，可知当  $y = \frac{\alpha}{\alpha+\beta}$  时，有最大值

$$M = \frac{B(a, b)}{B(a', b')} \left(\frac{\alpha}{\alpha+\beta}\right)^\alpha \left(\frac{\beta}{\alpha+\beta}\right)^\beta$$

通过计算可以知道， $\alpha, \beta$  越小时， $M$  越小。而为使  $a', b'$  都为整数，因此应当取  $a' = [a]$ ， $b' = [b]$

对 Gamma 分布，设  $V$  的参数为  $(a', b')$ ，则

$$\frac{f_Y(y)}{f_V(y)} = \frac{\Gamma(a')(b')^{a'}}{\Gamma(a)b^a} y^{a-a'} e^{-y(\frac{1}{b} - \frac{1}{b'})}$$

记  $\alpha = a - a'$ ， $\frac{1}{\beta} = \frac{1}{b} - \frac{1}{b'}$ ，则可以知道

当  $y = \alpha\beta$  时，取得最大值  $M = \frac{\Gamma(a')(b')^{a'}}{\Gamma(a)b^a} (\alpha\beta)^\alpha e^{-\alpha}$ ，为求得该式的最小值。

对  $a'$  来说， $a'$  越小，对应的极小值越小，因此应当取  $a' = [a]$ ，但是对于  $b'$ ，无法解析的给出取最小对应的值，只有具体情况数值求解。

**5.64** (a) 记  $A = \text{supp} f = \text{supp} g$ ，则

$$\begin{aligned} E\left(\frac{1}{m} \sum_{i=1}^m \frac{f(Y_i)}{g(Y_i)} h(Y_i)\right) &= E \frac{f(Y_i)}{g(Y_i)} h(Y_i) \\ &= \int_A \frac{f(y)}{g(y)} h(y) g(y) dy \\ &= \int_A f(y) h(y) dy \\ &= Eh(X) \end{aligned}$$

(b) 记  $T_i = \frac{f(Y_i)}{g(Y_i)} h(Y_i)$

$$E|T_i| = \int_A \left| \frac{f(y)}{g(y)} \right| |h(y) g(y)| dy = E|h(X)| \text{ 由大数定律:}$$

$$\frac{1}{m} \sum_{i=1}^m T_i \xrightarrow{P} ET_i = Eh(X)$$

(c) 记  $A_i = \frac{f(Y_i)}{g(Y_i)}, B_i = \frac{f(Y_i)}{g(Y_i)}h(Y_i)$

$EA_i = 1, E|A_i| = 1$  由大数定律  $\bar{A}_m \xrightarrow{p} 1$

$EB_i = h(X), E|B_i| < \infty$  由大数定律  $\bar{B}_m \xrightarrow{p} Eh(X)$

因此  $\frac{\bar{B}_m}{\bar{A}_m} \xrightarrow{p} \frac{Eh(X)}{1} = Eh(X)$ , 即  $\sum_{i=1}^m \frac{f(Y_i)/g(Y_i)}{\sum_{j=1}^m f(Y_j)/g(Y_j)} h(Y_i) \xrightarrow{p} Eh(X)$

(c) 当  $h$  是常数时,  $Eh(X) = h$ , (c) 给出的估计就为  $h = Eh(X)$ , 而 (b) 给出的估计量为  $Eh(X) \frac{1}{m} \sum_{i=1}^m \frac{f(Y_i)}{g(Y_i)}$ , 因此在这个特殊情况下, (c) 给出的估计量比 (b) 中公式更好。

### 模拟作业

[1] 由  $\int_0^2 g(x)dx = \frac{3}{2}$  故令  $K = \frac{2}{3}$ , 此时  $f(x) = Kg(x)$  为密度函数

[2]  $0 < x < 0.5$  时

$$F(x) = \int_0^x \frac{2}{3} 2x dx = \frac{2x^2}{3}$$

$0.5 < x < 1$  时

$$F(x) = \frac{1}{6} + \int_{0.5}^x \frac{2}{3} (2 - 2x) dx = \frac{1}{3} - \frac{2}{3} (x - 1)^2$$

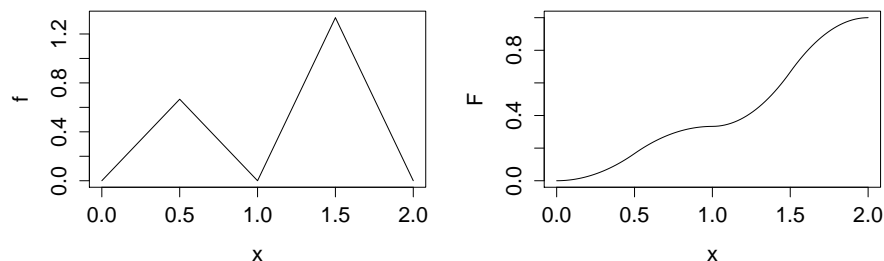
$1 < x < 1.5$  时

$$F(x) = \frac{1}{3} + \int_0^{x-1} \frac{2}{3} 4x dx = \frac{1}{3} + \frac{4}{3} (x - 1)^2$$

$1.5 < x < 2$  时

$$F(x) = \frac{2}{3} + \int_{1.5}^x \frac{2}{3} (8 - 4x) dx = 1 - \frac{4}{3} (x - 2)^2$$

做出  $f(x), F(x)$  的图形如下:



$$\text{计算 } \mu = EX = \frac{7}{6}, \sigma^2 = VarX = \frac{19}{72}$$

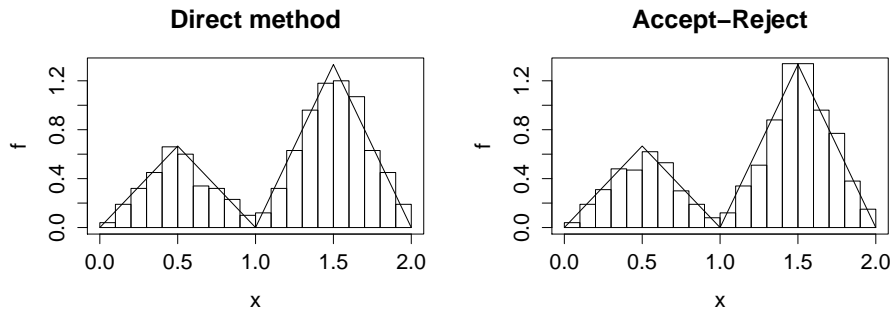
[3] 利用直接方法产生随机数:

用分段函数表示  $F^{-1}(x)$ , 作用于产生的 1000 个 (0,1) 均匀分布随即变量, 得到的 1000 个数即满足  $f(x)$  的分布, 计算所得的  $\hat{\mu}_1 = 1.179$

[4] 利用接收-拒绝法产生随机数:

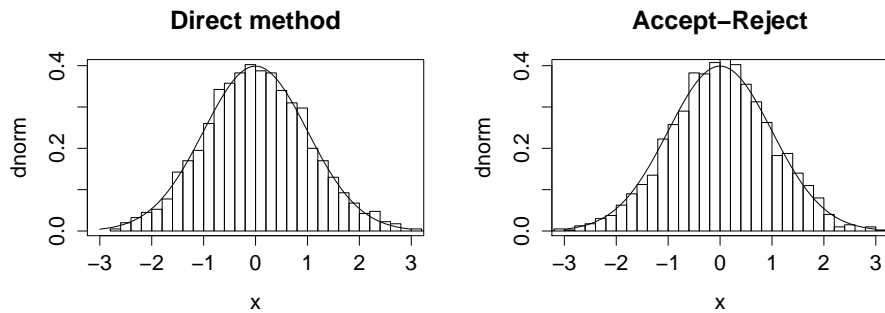
这里我为了避免使用循环语句, 而尽可能向量化, 采取了先生成两列足够多的均匀分布随机数列, 用向量操作筛选出满足  $y < f(x)$  数据的方法, 但是由于向量长度是固定的, 还要提前取定足够的数据使得筛选后至少有 1000 个数。这可以按书中的方法估计得到  $EN = \frac{8}{3}$ , 要得到 1000 个数据, 为了保证不会出现问题, 我先生成了 2800 组随即数。这样做虽然会额外生成一些数, 但是与直接作循环相比更能节省时间。

两种方法获得直方图与概率密度函数叠加所得图如下



[5] 两种方法各重复 2000 次：由于每次重复的步骤都是一样的，我没有写成循环，而是直接产生  $1000 \times 2000$  组数据，再分成 2000 组，分组求平均，这样与循环 2000 次是一样的，速度也要快一点。由大数定律，这 2000 个值近似服从正态分布，即  $\frac{\sqrt{1000}(\hat{\mu} - \mu)}{\sigma} \sim n(0, 1)$ ，可以做图来验证这一点，即这两千个数的直方图应当和正态分布密度函数相合。

两种方法得到的图像如下



两种方法得到的样本均值与方差为：

Direct method :  $\bar{\mu}_1 = 1.16631$ ,  $S_1^2 = 2.57 \times 10^{-4}$

Accept Reject :  $\bar{\mu}_2 = 1.16673$ ,  $S_2^2 = 2.56 \times 10^{-4}$

[6] 记时两种算法运行时间：都产生 2000 组平均值，直接方法用时 30.401s，接受拒绝法用时 75.579s，因此直接方法明显节约了很多时间。但这和问题的特殊性有关，因为 cdf 的反函数可以显示表达，在大多数情况下，cdf 并不能直接表达，而获取反函数需要解方程，而且这个方程一般会涉及积分，因此方程的求解会异常耗时。所以在反函数不能直接表达的时候，接受拒绝法要可行。