统计中的计算方法
第三次作业

于慧倩
14300180118

2017 年 5 月

1 第一题

用两种方法产生以下分布，并进行数值试验:

$$PX = i = \frac{e^{-\lambda}\lambda^i/i!}{\sum_{j=0}^{k} e^{-\lambda}\lambda^j/j!}, i = 0, \ldots, k$$

1.1 method 1: the inverse transform method

1.1.1 计算方法

生成在 $[0, 1]$ 均匀分布的随机数 U，继而生成:

$$X = \begin{cases} x_0 & \text{if } U < p_0 \\ x_1 & \text{if } p_0 \leqslant U < p_0 + p_1 \\ \cdots & \cdots \\ x_k & \text{if } \sum_{i=0}^{k-1} p_i \leqslant U < \sum_{i=0}^{k} p_i \end{cases}$$

通过下述方法生成 X

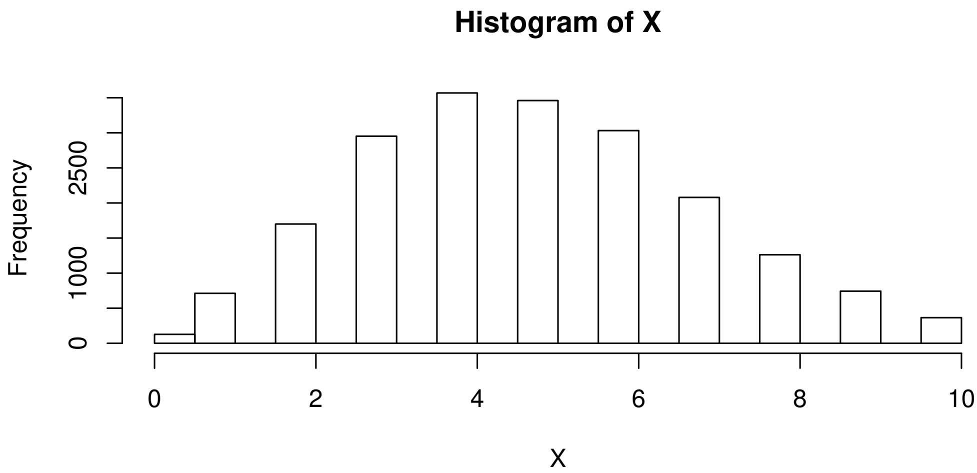
- 生成随机数 U
- 如果 $U < p_0$，令 $X = x_0$，停止。
- 否则，如果 $U < p_0 + p_1$，令 $X = x_1$，停止。
- ...
- 否则，如果 $U < \sum_{i=0}^{k-1}$，令 $X = x_{k-1}$，停止。
- 否则，令 $X = x_k$，停止。

由于 $p_i = P\{x = i\} = e^{-\lambda}\frac{\lambda^i}{i!}/\sum_{j=0}^{k} e^{-\lambda}\frac{\lambda^j}{j!}$，我们有 $p_{i+1} = p_i \frac{\lambda}{i+1}, i \geqslant 0$，所以实际上操作如下

- Step 1: 计算 $C = \sum_{j=0}^{k} e^{-\lambda}\frac{\lambda^j}{j!}$
- Step 2: 生成随机数 U
- Step 3: $i = 0, p = e^{-\lambda}/C, F = p$
- Step 4: 如果 $U \leqslant F$，令 $X = i$，停止
- Step 5: $i = i + 1; p = \lambda p/i; F = F + p;$
- Step 6: 重新回到 Step 4 继续判断

1.1.2 数值试验

取 $k=10,\lambda=5$ 为例进行数值试验，得到结果如下图所示。重复操作 20000 次得到长度为 20000 的分布如题干所示的随机数 X，且均值、方差分别为 $4.8859, 4.435303$



1.2 method 2:the acceptance-rejection method

1.2.1 计算方法

取 Y 为平均分布，即其概率密度函数 $\{q_j=\frac{1}{k}\}, j\geqslant 0$，令 $c=max(\frac{p_j}{q_j})$，计算可得 $c=\frac{k}{C}[\lambda-1]$。

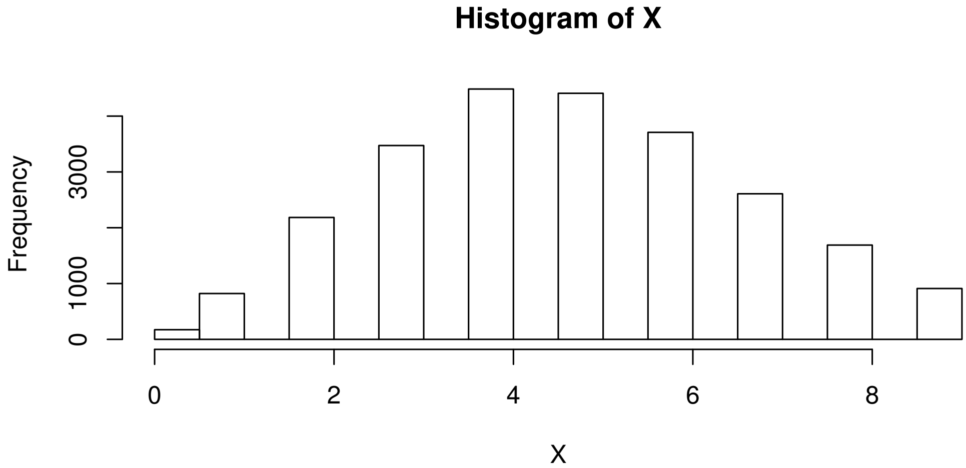
通过下述方法生成 X

- Step 1: 生成随机数 U_1，令 $Y=[kU_1]$
- Step 2: 生成随机数 U_2
- Step 3: 如果 $U_2\leqslant\frac{p_Y*k}{c}$，令 $X=Y$，停止。否则回到 Step 1.

1.2.2 数值试验

仍然取 $k=10,\lambda=5$ 为例进行数值试验，得到结果如下图所示。重复操作 10 万次得到长度为 24462 的分布如题干所示的随机数 X，且均值、方

差分别为 $4.816205, 4.03032$。并且 $\frac{1}{c} = 0.02465762$，实际有效操作比例为 0.024462



2 第二题

给出模拟下述分布的方法，并进行数值试验：

$$p_j = \begin{cases} 0.11 & \text{j is odd , and } 5 \leqslant j \leqslant 13 \\ 0.09 & \text{j is even, and } 6 \leqslant j \leqslant 14 \end{cases}$$

2.1 method 1:the inverse transform method

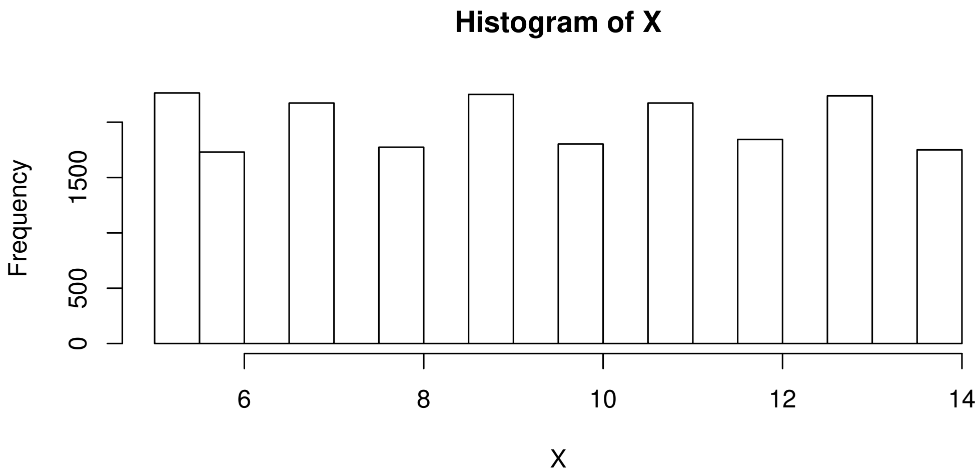
2.1.1 计算方法

通过下述方法生成 X

- Step 1: 生成随机数 U
- Step 2: $p = 0.11$, 令 $x = x_0 = 5$, 如果 $U < p$，令 $X = x$，停止。
- Step 3: 否则，$p = p + 0.09$,$x = x + 1$, 如果 $U < p$，令 $X = x$，停止。
- Step 4: 否则，$p = p + 0.11$,$x = x + 1$, 如果 $U < p$，令 $X = x$，停止。
- Step 5: 否则，回到 Step 3。

2.1.2 数值试验

进行数值试验，得到结果如下图所示。重复操作 2 万次得到长度为 2 万的分布如题干所示的随机数 X，且均值、方差分别为 9.45085 8.242596。



2.2 method 2:the acceptance-rejection method

2.2.1 计算方法

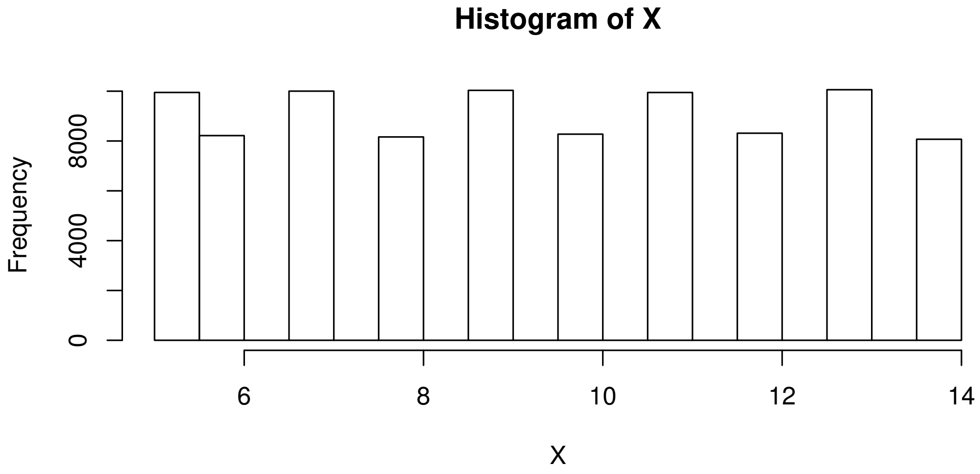
取 Y 为平均分布，即其概率密度函数 $\{q_j = \frac{1}{10}\}, j \geqslant 0$，令 $c = max(\frac{p_j}{q_j}) = 1.1$。

通过下述方法生成 X

- Step 1: 生成随机数 U_1，令 $Y = [10U_1] + 5$
- Step 2: 生成随机数 U_2
- Step 3: 如果 $U_2 \leqslant \frac{p_Y}{0.11}$，令 $X = Y$，停止。否则回到 Step 1.

2.2.2 数值试验

进行数值试验，得到结果如下图所示。重复操作 10 万次得到长度为 91020 的分布如题干所示的随机数 X，且均值、方差分别为 9.451362 8.221225。并且 $\frac{1}{c} = 0.9090909$，实际有效操作比例为 0.9102



3 第三题

给出具有如下概率密度函数的随机变量的产生方法，并进行数值验证：

$$f(x) = \begin{cases} e^{2x} & -\infty < x < 0 \\ e^{-2x} & 0 < x < \infty \end{cases}$$

3.1 the inverse transform algorithm

3.1.1 计算方法

因为 U 为 $(0,1)$ 上均匀分布，当 $x<0$ 时，有 $\int_{-\infty}^{x} e^{2t}\mathrm{d}t = U$, 得到 $x = \frac{1}{2}\ln(2U)$。当 $x \geqslant 0$ 时，有 $\int_{-\infty}^{0} e^{2t}\mathrm{d}t + \int_{0}^{x} e^{2t}\mathrm{d}t = U$, 得到 $x = -\frac{1}{2}\ln(2(1-U))$

，按照下述方法生成 X:

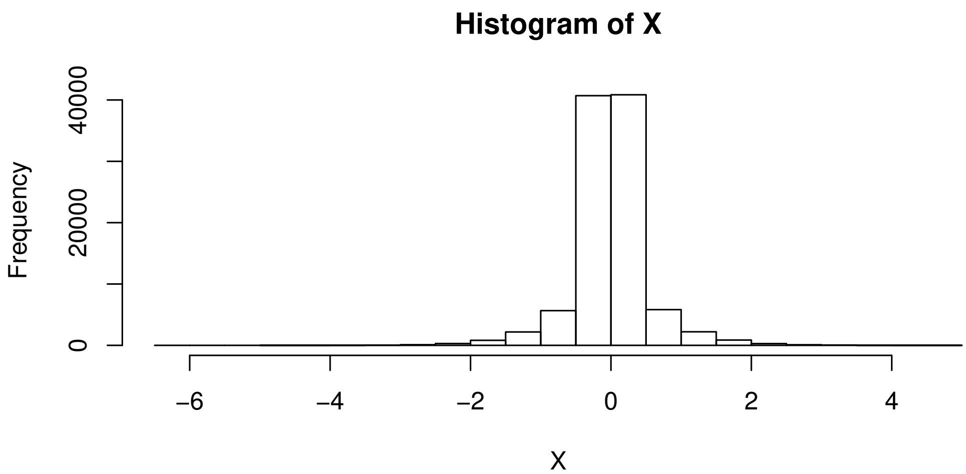
- Step 1: 生成 $(0,1)$ 上均匀分布随机数 U_1
- Step 3: 生成 $(0,1)$ 上均匀分布随机数 U_2
- Step 4: if $U_2 \leqslant \frac{1}{2}$，令 $X = \frac{1}{2}\ln(2U_1)$，否则 $X = -\frac{1}{2}\ln(2(1-U_1))$

3.1.2 数值试验

进行数值试验，得到结果如下图所示。

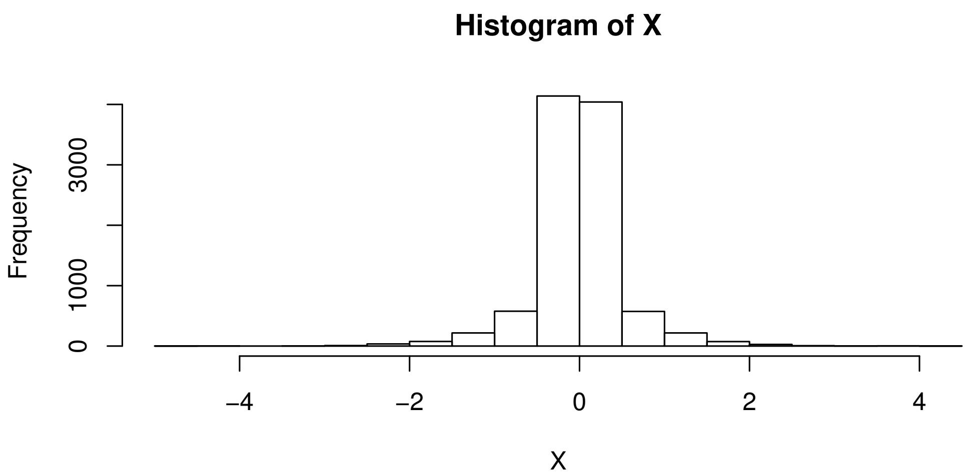
a) 重复操作 10 万次

得到长度为 10 万的分布如题干所示的随机数 X，且均值、方差分别为 $0.001559111, 0.2755747$。



b) 重复操作 1 万次

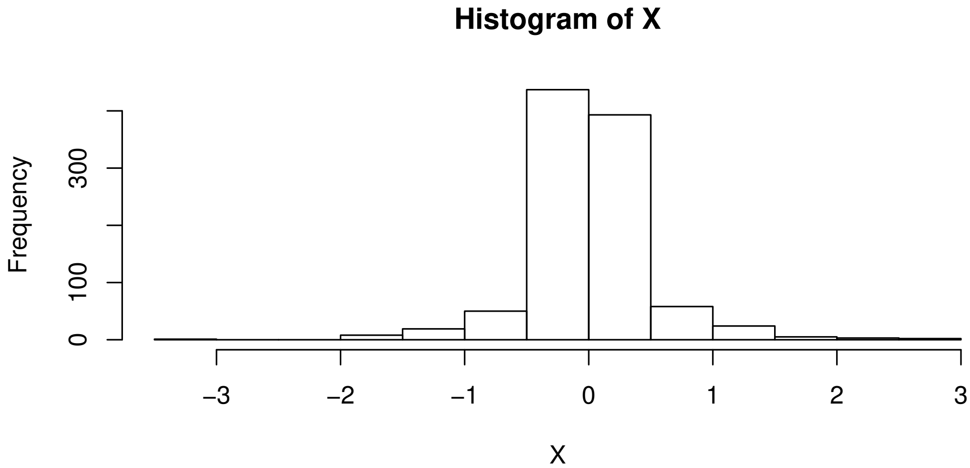
得到长度为 1 万的分布如题干所示的随机数 X，且均值、方差分别为 $-0.005648216, 0.2654377$。



c) 重复操作 1 千次

得到长度为 1 千的分布如题干所示的随机数 X，且均值、方差分别为

0.006721561, 0.2410665。



4 第四题

给出具有如下概率密度函数的随机变量的产生方法, 进行数值试验，并讨论方法的运算效率:$f(x) = 30(x^2 - 2x^3 + x^4), 0 \leqslant x \leqslant 1$.

4.1 method 1: the acceptance-rejection method

4.1.1 计算方法

令 Y 为均匀分布，其概率密度函数 $g(x) = 1, 0 < x < 1$，这样有 $c = max(\frac{f(x)}{g(x)}) = \frac{15}{8}$。

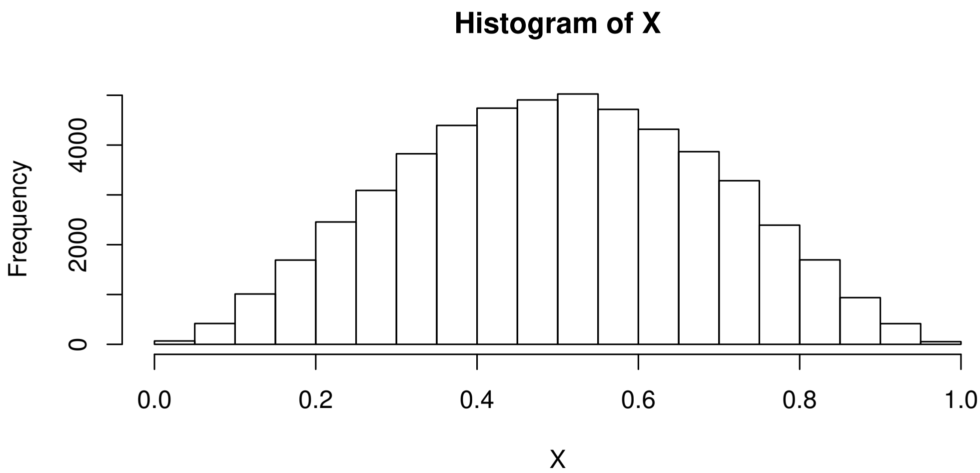
利用接受拒绝方法生成 X:

- Step 1: 生成 $(0,1)$ 上的均匀分布随机数 U_1，令 $Y = U_1$
- Step 2: 生成 $(0,1)$ 上的均匀分布随机数 U_2
- Step 3: 如果 $U_2 \leqslant 16Y^2(Y-1)^2$，令 $X = Y$，否则返回 Step 1.

4.1.2 数值试验

得到结果如下图所示, 重复操作 10 万次, 得到长度为 53298 的分布如题干所示的随机数 X，且均值、方差分别为 0.4998224, 0.03593751。并且

$\frac{1}{c} = 0.53333$，实际运算效率为 0.53298



4.2 method 2: the inverse transform method

4.2.1 计算方法

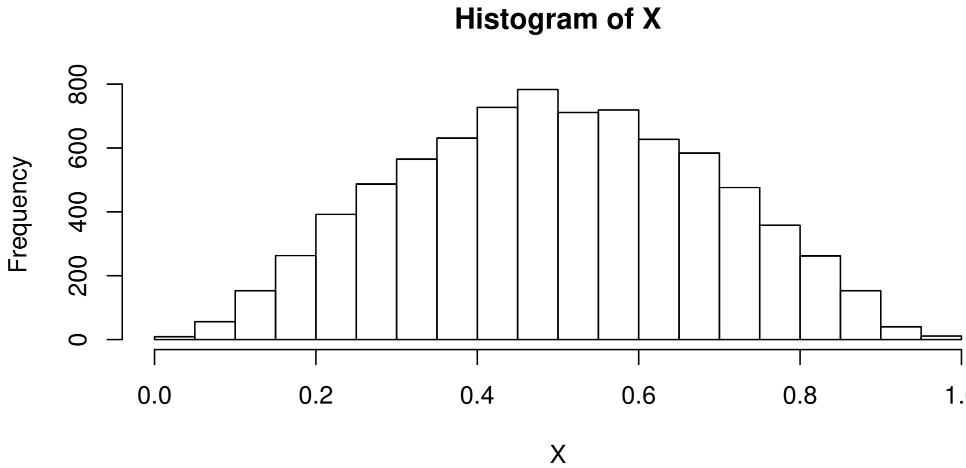
令 Y 为均匀分布，其概率密度函数 $g(x) = -6x(x-1), 0 < x < 1$，这样有 $c = max(\frac{f(x)}{g(x)}) = \frac{5}{4}$。

利用接受拒绝方法生成 X:

- Step 1: 生成 $(0,1)$ 上的均匀分布随机数 U_1
- Step 2: 求解 $2Y^3 + 3Y^2 = U_1$, 生成 Y;
- Step 3: 生成 $(0,1)$ 上的均匀分布随机数 U_2
- Step 4: 如果 $U_2 \leqslant -4Y(Y-1)$, 令 $X = Y$，否则返回 Step1.

4.2.2 数值试验

得到结果如下图所示, 重复操作 1 万次, 得到长度为 8007 的分布如题干所示的随机数 X, 且均值、方差分别为 $0.4979851, 0.03592379$。并且 $\frac{1}{c} = 0.8$，实际运算效率为 0.8007



5 第五题

给出产生具有如下概率密度函数的随机变量的接受拒绝方法，$f(x) = \frac{1}{2}x^2e^{-x}, 0 \leqslant x < \infty$. 假设用指数分布来产生此分布，给出最优的参数 λ.

5.1 计算最优参数

令 Y 为指数分布，其概率密度函数为 $g(x) = \lambda e^{-\lambda x}$，令 $c(x,\lambda) = max(\frac{f(x)}{g(x)} = \frac{1}{2\lambda}x^2e^{-x+\lambda x})$。

对 $c(x,\lambda)$ 关于 x 求导:

$$[2x + (\lambda - 1)]e^{-x+\lambda x} = 0$$

得到

$$x = \frac{2}{1-\lambda}$$

带入 $c(x,\lambda)$ 得到

$$c(\lambda) = \frac{1}{\lambda}\frac{2}{(1-\lambda)^2}e^{-2}$$

对 λ 求导令其为零，得到

$$\lambda = 1, \text{ or } , \lambda = \frac{1}{3}$$

最终有 $\lambda = \frac{1}{3}$ 为最优参数，此时 $c = \frac{27}{2}e^{-2}$

5.2 计算方法

按照 the acceptance-rejection method 生成 X:

- Step 1: 生成指数分布 Y，其概率密度函数为 $g(x)=\frac{1}{3}e^{-\frac{1}{3}x}$。
- Step 2: 生成 $(0,1)$ 上的均匀分布随机数，U
- Step 3: 如果 $U \leqslant \frac{1}{9}Y^2e^{2-\frac{2}{3}Y}$，令 $X=Y$，否则返回 Step 1.

其中利用 the inverse transform algorithm 生成指数分布 Y 步骤如下：

- Step 1: 生成 $(0,1)$ 上均匀分布随机数 U
- Step 2: 令 $Y=-3\ln(1-U)$。

5.3 数值试验

得到结果如下图所示, 重复操作 10 万次, 得到长度为 18094 的分布如题干所示的随机数 X，且均值、方差分别为 $3.007682, 2.975097$。并且 $\frac{1}{c}=0.5473375$，实际有效操作比例为 0.54957