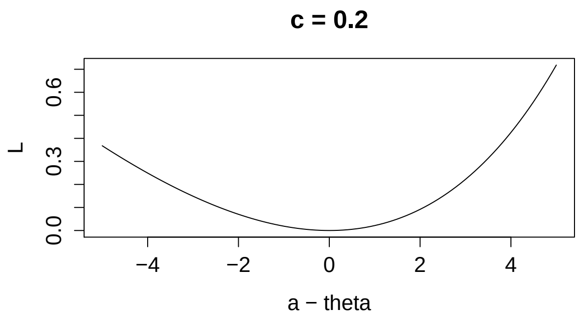
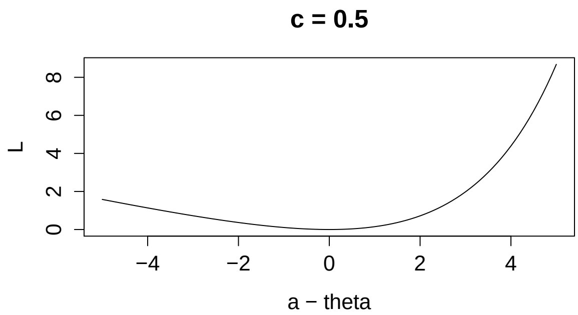
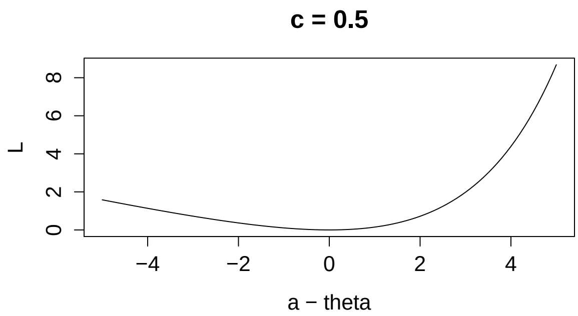
1 损失函数作业

7.65

(a) 画出函数图像如下:







可见随着 c 的增加，函数图像逐渐由对称变得不对称。

(b) $E(L(\theta,\delta)|\mathbf{x}) = \mathrm{e}^{c\delta}E(\mathrm{e}^{-c\theta}|\mathbf{x}) - c(\delta - E(\theta|\mathbf{x})) - 1$

求导可得$c\mathrm{e}^{c\delta}E(\mathrm{e}^{-c\theta}|\mathbf{x}) - c = 0 \Rightarrow \delta = -\frac{1}{c}\log E(\mathrm{e}^{-c\theta}|\mathbf{x})$

而$ce^{c\delta}E(e^{-c\theta}|\mathbf{x})-c$ 随δ 的单调递增，故极值为极小值, 满足要求，即：$\delta^{\pi}(x)=-\frac{1}{c}\log E(e^{-c\theta}|\mathbf{x})$

(c) 这里的$\pi(\theta)$ 代表全实轴的均匀分布函数，可先设$\pi(\theta)=\frac{1}{D}$ 最后令$D\to\infty$(其实后面的计算与 D 无关)

$f(\mathbf{x},\theta)=Ce^{-\sum_{i=1}^{n}\frac{(x_i-\bar{x})^2}{2\sigma^2}}e^{-\frac{n(\bar{x}-\theta)^2}{2\sigma^2}}=p(\mathbf{x})e^{-\frac{n(\bar{x}-\theta)^2}{2\sigma^2}}$

$m(\mathbf{x})=\int f(\mathbf{x}|\theta)\pi(\theta)d\theta=p(\mathbf{x})\int e^{-\frac{n(\bar{x}-\theta)^2}{2\sigma^2}}d\theta$

$\pi(\theta|\mathbf{x})=\dfrac{f(\mathbf{x},\theta)}{m(\mathbf{x})}=\dfrac{e^{-\frac{n(\bar{x}-\theta)^2}{2\sigma^2}}}{\int e^{-\frac{n(\bar{x}-\theta)^2}{2\sigma^2}}d\theta}$

即$\pi(\theta|\mathbf{x})\sim \mathrm{n}(\bar{x},\frac{\sigma^2}{n})$ 可以得到

$E(e^{-c\theta}|\mathbf{x})=M_{\theta|\mathbf{x}}(-c)=e^{-c\bar{x}+\frac{\sigma^2c^2}{2n}}$

$\delta^{B}(\bar{X})=\bar{X}-\dfrac{c\sigma^2}{2n}$

(d) $E(\theta|\mathbf{x})=\bar{x}$

$E(L(\theta,\delta^B)|\mathbf{x})=e^{c\delta}E(e^{-c\theta}|\mathbf{x})-c(\delta-E(\theta|\mathbf{x}))-1=\dfrac{c^2\sigma^2}{2n}$

$E(L(\theta,\bar{x})|\mathbf{x})=e^{c\bar{x}}E(e^{-c\theta}|\mathbf{x})-c(\bar{x}-E(\theta|\mathbf{x}))-1=e^{\frac{\sigma^2c^2}{2n}}-1$

(e) $L_2(\theta,a)=|\theta-a|^2$

$E(L_2(\theta,\delta^B)|\mathbf{x})=E\left[(\bar{x}-\theta)^2+(\dfrac{c\sigma^2}{2n})^2-2(\bar{x}-\theta)\dfrac{c\sigma^2}{2n}\Big|x\right]=\dfrac{\sigma^2}{n}+\left(\dfrac{c\sigma^2}{2n}\right)^2$

$E(L_2(\theta,\bar{x})|\mathbf{x})=E\left[(\bar{x}-\theta)^2\Big|x\right]=\dfrac{\sigma^2}{n}$

7.66

(a) $\because\ n\bar{X}\sim \mathrm{B}(n,\theta)\mathrm{Var}(n\bar{X})=E(n\bar{X})^2-(En\bar{X})^2\neq\theta^2$

(b) $\therefore E\bar{X}^2=\dfrac{1}{n^2}E(n\bar{X})^2=\dfrac{1}{n^2}\left[n\theta(1-\theta)+n^2\theta^2\right]=\theta^2+\dfrac{\theta(1-\theta)}{n}$

$T_n=\bar{X}^2$

$T_n^{(i)}=\left(\dfrac{\sum_{k=1}^{n}X_k-X_i}{n-1}\right)^2=\left(\dfrac{\sum_{k=1}^{n}X_k}{n-1}\right)^2+\left(\dfrac{X_i}{n-1}\right)^2-\dfrac{2X_i(\sum_{k=1}^{n}X_k)}{(n-1)^2}$

$$\sum_{i=1}^{n} T_n^{(i)} = \frac{(n-2)(\sum_{k=1}^n X_k)^2 + \sum_{i=1}^n X_i^2}{(n-1)^2} = \frac{(n-2)n^2 T_n + \sum_{i=1}^n X_i^2}{(n-1)^2}$$

$$\begin{aligned} JK(T_n)1 &= nT_n - \frac{n-1}{n}\sum_{i=1}^{n} T_n^{(i)} \\ &= nT_n - \frac{n(n-2)T_n}{n-1} - \frac{\sum_{i=1}^n X_i^2}{n(n-1)} \\ &= \frac{n}{n-1}T_n - \frac{\sum_{i=1}^n X_i^2}{n(n-1)} \end{aligned}$$

而对 0-1 分布$X_i = X_i^2$, 故$JK(T_n) = \dfrac{n}{n-1}T_n - \dfrac{\bar{X}}{(n-1)}$

(c) $EJK(T_n) = \frac{n}{n-1}(\theta^2 + \frac{\theta(1-\theta)}{n}) - \frac{\theta}{n-1} = \theta$

(d) 伯努利分布属于指数分布族，易知$\bar{X}$ 为θ 的完全充分统计量。

而$JK(T_n) = \dfrac{n}{n-1}\bar{X}^2 + \dfrac{\bar{X}}{n-1}$ 是$\bar{X}$ 的函数，因此$JK(T_n)$ 是θ^2 的最佳无偏估计

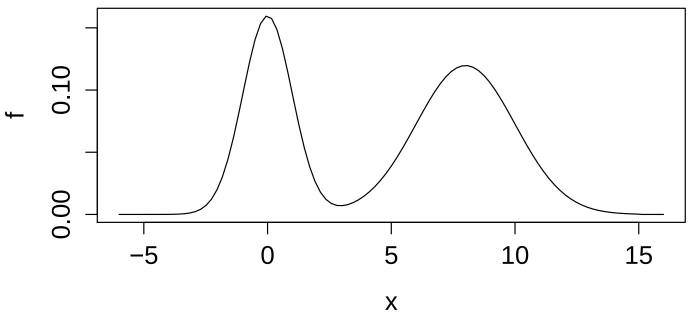
随机试验验证 利用 R 生成随机数，选取$\theta = 0.3, 0.5, 0.7$，分别生成 10 个随机数 (n 不能选的太大，否则$\frac{\theta(1-\theta)}{n}$ 趋近 0 而看不出区别)，计算估计量T_n, $JK(T_n)$，并重复 10000 次，再计算这 10000 次的平均值，可以验证上题的 (a),(c) 两小问。

mean	$\theta = 0.3$	$\theta = 0.5$	$\theta = 0.7$
θ^2	0.09	0.25	0.49
T_n	0.11179	0.27567	0.51027
JKT_n	0.09072	0.25066	0.48928

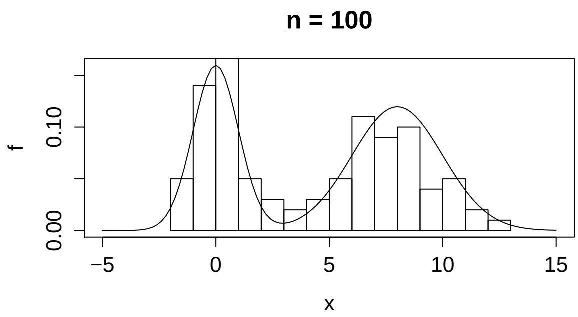
T_n 的平均值明显比θ^2 要大，而JKT_n 的平均值与θ^2 很接近。

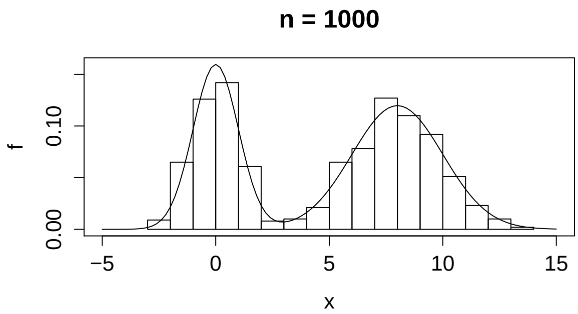
2 随机模拟：密度函数的估计

1 求出 C 的值为$C = 1.0001397$, f 的函数图像如下：



2 将密度函数与直方图画在一张图上，可以看到抽样结果应当是正确的:

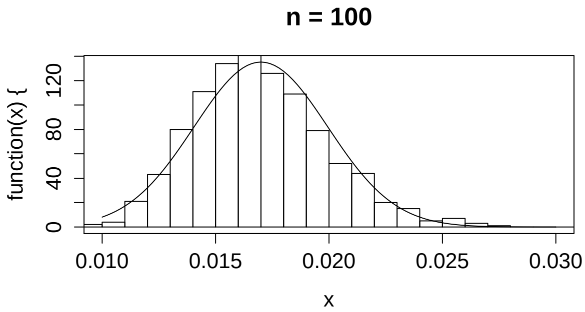


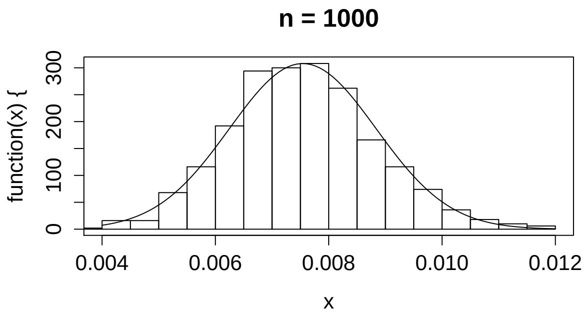


3 计算出 1000 次模拟所得的 RAISE 值得均值和方差为如下表所示

RAISE	mean	sd
n = 100	0.01691	0.00301
n = 1000	0.00750	0.00129

而且可以发现 RAISE 的分布基本服从正态分布 (证明显得有些困难)。直方图如下





可以看出取样数量 n 越大，密度函数的估计越好。这从直观上也容易理解，样本越多，对应取得的可以取得的 h 越小，密度函数的估计越接近真实的密度函数。