概率论与数理统计 B　第十一，十二周作业　5 月 9 日 周六

PB18151866 龚小航

7.40. 设 $X_1,\dots,X_n$ 为抽自指数分布 $f(x,\mu)=e^{-(x-\mu)},\quad x\ge\mu,\quad -\infty<\mu<+\infty$ 的简单样本.

(1) 试求 μ 的极大似然估计 $\hat{\mu}^*$，$\hat{\mu}^*$ 是 μ 的无偏估计吗?

如果不是，试对它作修改，以得到 μ 的无偏估计 $\hat{\mu}^{**}$.

(2) 试求 μ 的矩估计 $\hat{\mu}$，并证明它是 μ 的无偏估计.

(3) 试问 $\hat{\mu}^{**}$ 和 $\hat{\mu}$ 哪一个有效?

解：(1) 显然 $f(x;\mu)$ 关于 μ 是单调的，似然函数为：

$$L(\mu)=\prod_{i=1}^{n}f(x_i;\mu)=\prod_{i=1}^{n}e^{-(x_i-\mu)}=e^{-\sum_{i=1}^{n}(x_i-\mu)}==e^{n\mu-\sum_{i=1}^{n}x_i}$$

它关于 μ 单调递增。因此 μ 要取最大值。由于有约束 $x\ge\mu$，因此 μ 最大能取 $X_{(1)}$，即抽出样本的最小元素。

$$\Longrightarrow\quad \hat{\mu}^*=X_{(1)}$$

再说明 $\hat{\mu}^*$ 是否为 μ 的无偏估计：

由于 $X_{(1)}$ 是抽出样本的最小值，因此计算它的均值时 $E\left(X_{(1)}\right)\neq E(X)$，必须通过分布函数来计算。

$$F_{X_{(1)}}(x)=1-P(X_i>x)=1-\prod_{i=1}^{n}\left(1-F_{X_i}(x)\right)=1-e^{-n(x-\mu)}$$

$$\text{即有}\quad f_{X_{(1)}}=F'_{X_{(1)}}(x)=ne^{-n(x-\mu)}$$

由此就可以通过均值的定义式来计算 $E(\hat{\mu}^*)$：

$$E(\hat{\mu}^*)=E\left(X_{(1)}\right)=\int_{\mu}^{\infty}xne^{-n(x-\mu)}\,dx=\left(-xe^{-n(x-\mu)}\Big|_{x=\mu}^{\infty}\right)+\int_{\mu}^{\infty}e^{-n(x-\mu)}\,dx=\mu+\frac{1}{n}$$

因此 $\hat{\mu}^*$ 不是 μ 的无偏估计，$\hat{\mu}^{**}=\hat{\mu}^*-\frac{1}{n}=X_{(1)}-\frac{1}{n}$

(2) 直接按定义求其矩估计：

$$E(x)=\int_{\mu}^{\infty}xe^{-(x-\mu)}\,dx=\left(-xe^{-(x-\mu)}\Big|_{x=\mu}^{\infty}\right)+\int_{\mu}^{\infty}e^{-(x-\mu)}\,dx=\mu+1$$

用样本均值估计总体均值，立刻可得：

$$\hat{\mu}=\overline{X}-1$$

再证明它是 μ 的无偏估计：

$$E(\hat{\mu})=E\left(\overline{X}-1\right)=E\left(\overline{X}\right)-1=\frac{1}{n}\sum_{i=1}^{n}E(X_i)-1=\frac{1}{n}\cdot n(\mu+1)-1=\mu$$

因此 $\hat{\mu}$ 就是 μ 的无偏估计。

(3) 均方误差小的估计更有效：

$$Var(\hat{\mu}^{**})=Var\left(X_{(1)}-\frac{1}{n}\right)=Var\left(X_{(1)}\right)=E\left(\left(X_{(1)}\right)^2\right)-E^2\left(X_{(1)}\right)$$

为此还先需要计算 $E\left(\left(X_{(1)}\right)^2\right)$。利用两次分部积分，立刻可得：

$$E\left(\left(X_{(1)}\right)^2\right)=\int_{\mu}^{\infty}x^2ne^{-n(x-\mu)}\,dx=\mu^2+\frac{2}{n}\mu+\frac{2}{n^2}$$

$$\Longrightarrow\quad Var(\hat{\mu}^{**})=\mu^2+\frac{2}{n}\mu+\frac{2}{n^2}-\left(\mu+\frac{1}{n}\right)^2=\frac{1}{n^2}$$

另一方面，有：

$$Var(\hat{\mu})=Var\left(\overline{X}-1\right)=Var\left(\overline{X}\right)=E\left(\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)^2\right)-E^2\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=1$$

显然有 $Var(\hat{\mu})=1>\frac{1}{n^2}=Var(\hat{\mu}^{**})$

因此 $\hat{\mu}^{**}$ 更有效。

7.46. 设X_1,X_2,X_3 i.i.d. 服从均匀分布 $U(0,\theta)$,试证$\frac{4}{3}\max\limits_{1\le i\le 3} X_i$ 及 $4\min\limits_{1\le i\le 3} X_i$ 都是θ的无偏估计量. 哪个更有效?

解：将X_1,X_2,X_3重新排列，得到由小到大的排列 $X_{(1)},X_{(2)},X_{(3)}$

因此题中的两个统计量分别可以表示为：

$$\frac{4}{3}X_{(3)}\ ;\ 4X_{(1)}$$

接下来就需要求其均值，为此先要求出分布函数，通过求导得到概率密度函数。

$$F_{X_{(3)}}(x)=P(X_1,X_2,X_3\le x)=\prod_{i=1}^{3}F_{X_i}(x)=\left(\frac{x}{\theta}\right)^3 \quad\Rightarrow\quad f_{X_{(3)}}(x)=F'_{X_{(3)}}(x)=\frac{3x^2}{\theta^3}$$

$$F_{X_{(1)}}(x)=1-P(X_1,X_2,X_3\ge x)=1-\prod_{i=1}^{3}\left(1-F_{X_i}(x)\right)=1-\left(1-\frac{x}{\theta}\right)^3 \quad\Rightarrow\quad f_{X_{(1)}}(x)=-\frac{3}{\theta}\left(1-\frac{x}{\theta}\right)^2$$

接下来就可以说明以上两个估计量都是θ的无偏估计量：

$$E\left(\frac{4}{3}X_{(3)}\right)=\frac{4}{3}E\left(X_{(3)}\right)=\frac{4}{3}\int_0^\theta x\frac{3x^2}{\theta^3}\ dx=\theta$$

$$E\left(4X_{(1)}\right)=4E\left(X_{(1)}\right)=4\int_0^\theta x\left(-\frac{3}{\theta}\left(1-\frac{x}{\theta}\right)^2\right)\ dx=\theta$$

所以这两个统计量都是θ的无偏估计。

接下来再比较哪个量的均方误差更小，小的统计量更有效：

为此先要计算$E\left(\left(\frac{4}{3}X_{(3)}\right)^2\right)$和$E\left(\left(4X_{(1)}\right)^2\right)$：

$$E\left(\left(\frac{4}{3}X_{(3)}\right)^2\right)=\frac{16}{9}\int_0^\theta x^2\frac{3x^2}{\theta^3}\ dx=\frac{16}{15}\theta^2$$

$$E\left(\left(4X_{(1)}\right)^2\right)=16\int_0^\theta x^2\left(-\frac{3}{\theta}\left(1-\frac{x}{\theta}\right)^2\right)\ dx=\frac{8}{5}\theta^2$$

因此，有：

$$Var\left(\frac{4}{3}X_{(3)}\right)=E\left(\left(\frac{4}{3}X_{(3)}\right)^2\right)-E^2\left(\frac{4}{3}X_{(3)}\right)=\frac{16}{15}\theta^2-\theta^2=\frac{1}{15}\theta^2$$

$$Var\left(4X_{(1)}\right)=E\left(\left(4X_{(1)}\right)^2\right)-E^2\left(4X_{(1)}\right)=\frac{8}{5}\theta^2-\theta^2=\frac{3}{5}\theta^2$$

显然有 $Var\left(\frac{4}{3}X_{(3)}\right)<Var\left(4X_{(1)}\right)$ 成立。

$$\therefore\quad \frac{4}{3}\max_{1\le i\le 3}X_i \quad \text{更有效}$$

7.63. 随机从一批钉子中抽取 9 枚，测得其长度（cm）为：

2.15, 2.13, 2.10, 2.14, 2.15, 2.16, 2.12, 2.11, 2.13,

假设钉子长度服从正态分布，分别在下面两种情况下，求出总体均值的 90% 置信区间：

(1) $\sigma=0.01$ (2) σ未知.

解：(1) 记取出样本的长度为随机变量X，先求出样本均值：

$$\overline{X}=\frac{1}{9}\sum_{i=1}^{9}x_i=\frac{1919}{900}\approx 2.1322\ \text{cm}$$

当 σ 已知时，可选取枢轴变量：

$$S=\frac{\sqrt{n}(\overline{X}-\mu)}{\sigma},\qquad \text{且有 } S\sim N(0,1)$$

要求置信系数 $1-\alpha=90\%$，$\alpha=0.10$ 写出置信区间：

$$[\hat{\mu}_1,\hat{\mu}_2]=[\ \overline{X}-\frac{\sigma}{\sqrt{n}}u_{\alpha/2}\ ,\ \overline{X}+\frac{\sigma}{\sqrt{n}}u_{\alpha/2}\ ]$$

其中$u_{\alpha/2}$满足条件 $P\left(|X|\ge u_{\alpha/2}\right)=\alpha$.通过查标准正态分布双侧上分位点表即可知：

$$u_{0.10/2}=1.6449$$

带入全部数据，即可知：

$$[\hat{\mu}_1,\hat{\mu}_2]=\left[\frac{1919}{900}-\frac{0.01}{\sqrt{9}}1.6449\ ,\ \frac{1919}{900}+\frac{0.01}{\sqrt{9}}1.6449\right]=[2.127\ ,\ 2.138]$$

(2) σ 未知时，使用上面的枢轴变量已经不可行，因为其中带有未知量σ.

选取新的枢轴变量T，令：

$$T=\frac{\sqrt{n}(\overline{X}-\mu)}{S}\ ,\quad \text{且 } T\sim t_{n-1}$$

要求置信系数 $1-\alpha=90\%$，$\alpha=0.10$ 写出置信区间：

$$[\hat{\mu}_1,\hat{\mu}_2]=\left[\overline{X}-\frac{S}{\sqrt{n}}t_{n-1}\left(\frac{\alpha}{2}\right)\ ,\ \overline{X}+\frac{S}{\sqrt{n}}t_{n-1}\left(\frac{\alpha}{2}\right)\right]$$

其中 $t_n(\alpha)$ 满足条件 $P\left(|X|\ge t_n(\alpha)\right)=\alpha$. 通过查 t 分布上侧分位点表即可知：

$$t_{n-1}\left(\frac{\alpha}{2}\right)=t_8(0.05)=1.860$$

再求出样本方差S：

$$S=\sqrt{\frac{1}{9-1}\sum_{i=1}^{9}\left(x_i-\overline{X}\right)^2}=\sqrt{\frac{71}{180000}}=0.01986$$

带入全部数据，即可知：

$$[\hat{\mu}_1,\hat{\mu}_2]=\left[\frac{1919}{900}-\frac{\sqrt{\frac{71}{180000}}}{\sqrt{9}}1.860\ ,\ \frac{1919}{900}+\frac{\sqrt{\frac{71}{180000}}}{\sqrt{9}}1.860\right]=[2.120\ ,\ 2.144]$$

7.67. 一家企业更换了领导，采取了新的经营策略. 随机选取公司 11 种商品，更换经营策略前后一个季度的销量（万元）如表，假设销量服从正态分布.

前	69.3	38.0	131.4	123.1	127.3	57.7	95.7	89.4	93.8	102.0	73.3
后	72.5	33.5	132.1	129.8	121.2	54.0	104.6	92.6	119.4	84.7	85.1

(1) 更换经营策略前平均销量的 95% 置信区间;

(2) 更换经营策略后平均销量的 95% 置信区间;

(3) 更换经营策略前后平均销量差异的 95% 置信区间.

解：销量服从正态分布，即服从分布 $N(\mu,\sigma^2)$.先对样本取良好的点估计，这里取均值$\overline{X}$

$$\overline{X_{前}}=\frac{1}{11}\sum x_i=91\ ;\quad \overline{X_{后}}=\frac{1}{11}\sum x_i=\frac{2059}{22}=93.5909$$

$$S_{前}=\sqrt{\frac{1}{11-1}\sum_{i=1}^{11}\left(x_i-\overline{X_{前}}\right)^2}=29.7251\ ;\quad S_{后}=\sqrt{\frac{1}{11-1}\sum_{i=1}^{11}\left(x_i-\overline{X_{后}}\right)^2}=31.8103$$

(1) 销量服从正态分布，需要估计其均值 μ. 由于 σ 为未知量，选取枢轴变量T:

$$T=\frac{\sqrt{n}(\overline{X_{前}}-\mu)}{S_{前}}\ ,\quad 且\ T\sim t_{n-1}$$

要求置信系数 $1-\alpha=95\%$, $\alpha=0.05$ 写出置信区间:

$$[\hat{\mu}_1,\hat{\mu}_2]=\left[\overline{X_{前}}-\frac{S_{前}}{\sqrt{n}}t_{n-1}\left(\frac{\alpha}{2}\right)\ ,\ \overline{X_{前}}+\frac{S_{前}}{\sqrt{n}}t_{n-1}\left(\frac{\alpha}{2}\right)\right]$$

其中 $t_n(\alpha)$ 满足条件 $P\big(|X|\geq t_n(\alpha)\big)=\alpha$. 通过查 t 分布上侧分位点表即可知:

$$t_{n-1}\left(\frac{\alpha}{2}\right)=t_{10}(0.025)=2.208$$

将所有已知量全部带入，即可得:

$$[\hat{\mu}_1,\hat{\mu}_2]=[\,71.211\ ,\ 110.789\,]$$

(2) 同上一问，选取枢轴变量T:

$$T=\frac{\sqrt{n}(\overline{X_{后}}-\mu)}{S_{后}}\ ,\quad 且\ T\sim t_{n-1}$$

要求置信系数 $1-\alpha=95\%$, $\alpha=0.05$ 写出置信区间:

$$[\hat{\mu}_1,\hat{\mu}_2]=\left[\overline{X_{后}}-\frac{S_{后}}{\sqrt{n}}t_{n-1}\left(\frac{\alpha}{2}\right)\ ,\ \overline{X_{后}}+\frac{S_{后}}{\sqrt{n}}t_{n-1}\left(\frac{\alpha}{2}\right)\right]$$

其中 $t_n(\alpha)$ 满足条件 $P\big(|X|\geq t_n(\alpha)\big)=\alpha$. 通过查 t 分布上侧分位点表即可知:

$$t_{n-1}\left(\frac{\alpha}{2}\right)=t_{10}(0.025)=2.208$$

将所有已知量全部带入，即可得:

$$[\hat{\mu}_1,\hat{\mu}_2]=[\,72.414\ ,\ 114.768\,]$$

(3) 两个正态总体需要求均值的差值分布，即需要求 $\mu_x-\mu_y$ 的95%区间估计。【教材 p178 例 4.2】

先计算S:

$$S=\frac{1}{\sqrt{n+m-2}}\sqrt{\sum_{i=1}^{n}\left(x_i-\overline{X}\right)^2+\sum_{i=1}^{m}\left(y_i-\overline{Y}\right)^2}=\frac{1}{\sqrt{20}}\sqrt{\frac{441791}{50}+10118.929}=30.785$$

此时选取枢轴变量T为:

$$T=\sqrt{\frac{mn}{m+n}}\frac{\left(\left(\overline{X}-\overline{Y}\right)-\left(\mu_x-\mu_y\right)\right)}{S}\ ,\ 且\ T\sim t_{n+m-2}$$

然后给出区间估计:

$$\left[\widehat{\mu_x-\mu_y}_1,\widehat{\mu_x-\mu_y}_2\right]=\left[\left(\overline{X}-\overline{Y}\right)-\sqrt{\frac{m+n}{mn}}St_{n+m-2}\left(\frac{\alpha}{2}\right)\ ,\ \left(\overline{X}-\overline{Y}\right)+\sqrt{\frac{m+n}{mn}}St_{n+m-2}\left(\frac{\alpha}{2}\right)\right]$$

要求置信系数 $1-\alpha=95\%$, $\alpha=0.05$; $m=n=11$

其中 $t_n(\alpha)$ 满足条件 $P\big(|X|\geq t_n(\alpha)\big)=\alpha$. 通过查 t 分布上侧分位点表即可知:

$$t_{n+m-2}\left(\frac{\alpha}{2}\right)=t_{20}(0.025)=2.086$$

将所有已知量全部带入，即可得:

$$\left[\widehat{\mu_x-\mu_y}_1,\widehat{\mu_x-\mu_y}_2\right]=[-29.974\ ,\ 24.792\,]$$

7.71. 一批零件的长度 $X \sim N(\mu,\sigma^2)$, 从这批零件中随机地抽 10 件, 测得长度值分别为 (单位: mm):

49.5， 50.4， 49.7， 51.1， 49.4， 49.7， 50.8， 49.9， 50.3， 50.0.

在下列条件下求这批零件长度总体方差σ^2的 95% 置信区间.

(1) $\mu = 50$mm. (2) μ 未知. 【教材 p180 例 4.3】

解: (1) 单个正态分布已知μ来估计σ^2,由教材第二章 92 页给出的关系，可令枢轴变量T为:

$$T=\frac{(n-1)S^2}{\sigma^2}=\frac{\sum_{i=1}^{n}\left(x_i-\overline{X}\right)^2}{\sigma^2}=\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{\sigma^2}, \quad 且\ T\sim\chi_{n-1}^2$$

要求置信系数 $1-\alpha=95\%$，$\alpha=0.05$；

由此可以写出置信区间:

$$[\hat{\sigma}_1^2,\hat{\sigma}_2^2]=\left[\frac{(n-1)S^2}{\chi_{n-1}^2\left(\frac{\alpha}{2}\right)},\ \frac{(n-1)S^2}{\chi_{n-1}^2\left(1-\frac{\alpha}{2}\right)}\right]=\left[\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{\chi_{n-1}^2\left(\frac{\alpha}{2}\right)},\ \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{\chi_{n-1}^2\left(1-\frac{\alpha}{2}\right)}\right]$$

再计算各种未知量。由于样本已知，可求得$(n-1)S^2$：代入数据计算，可得

$$\sum_{i=1}^{n}(x_i-\mu)^2=2.9$$

要求置信系数 $1-\alpha=95\%$，$\alpha=0.05$

其中 $\chi_n^2(\alpha)$ 满足条件 $P\left(|X|>\chi_n^2(\alpha)\right)=\alpha$. 通过查卡方分布上侧分位点表即可知:

$$\chi_{n-1}^2\left(\frac{\alpha}{2}\right)=\chi_9^2(0.025)=19.023\ ; \qquad \chi_{n-1}^2\left(1-\frac{\alpha}{2}\right)=\chi_9^2(0.975)=2.700$$

将所有已知量全部带入，即可得:

$$[\hat{\sigma}_1^2,\hat{\sigma}_2^2]=[0.1524，1.0741]$$

(2) μ未知时，由教材第二章 92 页给出的关系，可令枢轴变量T为:

$$T=\frac{(n-1)S^2}{\sigma^2}, \quad 且\ T\sim\chi_{n-1}^2$$

其中，计算样本方差S^2:

$$\overline{X}=\frac{1}{n}\sum_{i=1}^{10}x_i=50.08\ ; \quad S^2=\frac{1}{10-1}\sum_{i=1}^{10}\left(x_i-\overline{X}\right)^2=\frac{709}{2250}=0.3151$$

要求置信系数 $1-\alpha=95\%$，$\alpha=0.05$；

由此可以写出置信区间: $\chi_{n-1}^2\left(\frac{\alpha}{2}\right)=\chi_9^2(0.025)=19.023$； $\chi_{n-1}^2\left(1-\frac{\alpha}{2}\right)=\chi_9^2(0.975)=2.700$

$$[\hat{\sigma}_1^2,\hat{\sigma}_2^2]=\left[\frac{(n-1)S^2}{\chi_{n-1}^2\left(\frac{\alpha}{2}\right)},\ \frac{(n-1)S^2}{\chi_{n-1}^2\left(1-\frac{\alpha}{2}\right)}\right]=[0.1491，1.0504]$$

7.76. 假设到一商场的顾客有 p 的概率购买商品，现随机抽取了 500 个顾客，其中 15 个购买了商品. 求 p 的 95% 置信区间.

解: 一个顾客是否购买商品只有两种取值，买$(x=1)$或者不买$(x=0)$，因此是一个二项分布

记购买商品的顾客人数为随机变量X_n. 由于商场的顾客人数比较多，可以利用中心极限定理得到购买人数的分布规律:

$$\lim_{n\to\infty}\frac{X_n-np}{\sqrt{np(1-p)}}\sim N(0,1)$$

取枢轴变量T:

$$T=\frac{X_n-np}{\sqrt{np(1-p)}}$$

则有:

$$P\left(-u_{\alpha/2}\le T\le u_{\alpha/2}\right)\approx 1-\alpha$$

从上述不等式中反解p，即解 $A\le p\le B$

经过计算，可以解出:

$$A=\frac{n}{n+u_{\alpha/2}^2}\left(\frac{X_n}{n}+\frac{u_{\alpha/2}^2}{2n}+u_{\alpha/2}\sqrt{\frac{\frac{X_n}{n}\left(1-\frac{X_n}{n}\right)}{n}+\frac{u_{\alpha/2}^2}{4n^2}}\right)=0.018$$

$$B=\frac{n}{n+u_{\alpha/2}^2}\left(\frac{X_n}{n}+\frac{u_{\alpha/2}^2}{2n}-u_{\alpha/2}\sqrt{\frac{\frac{X_n}{n}\left(1-\frac{X_n}{n}\right)}{n}+\frac{u_{\alpha/2}^2}{4n^2}}\right)=0.049$$

由此，可知 p 的95%置信区间为:

$$[\hat{p}_1,\hat{p}_2]=[0.018，0.049]$$

7.81. 设一农作的单位面积产量服从正态分布 $N(80,\sigma^2)$，其标准差$\sigma=5$，问至少需要几块试验田，才能有99%的把握保证这些试验田的单位面积平均产量大于 75？

解：要求置信系数 $1-\alpha=99\%$，$\alpha=0.01$；$\mu=80,\sigma=5$

记一块试验田的单位产量为随机变量 X

选取枢轴变量T：

$$T=\frac{\sqrt{n}(\overline{X}-\mu)}{\sigma}\text{，且 } T\sim N(0,1)$$

$$P(T\geq -u_\alpha=-u_{0.01})=1-\alpha$$

利用题中所给的不等式条件，均值大于 75 即可转化为 n 的不等式：

$$\overline{X}\geq \mu+u_\alpha\frac{\sigma}{\sqrt{n}}\geq 75 \quad\Longrightarrow\quad n\geq\left(\frac{u_{0.02}\sigma}{\mu-75}\right)^2=5.4140$$

因此需要六块试验田即可。

8.1. 假设 $X_1\ldots\ldots,X_{16}$ 服从正态分布 $N(\mu,0.16)$.

检验问题 $H0$：$\mu=0.5$ $\leftrightarrow$ $H1$：$\mu>0.5$ 显著水平为 0.05.

(1) 检验的拒绝域是什么？ (2) $\mu=0.65$ 时犯第二类错误的概率是多少？

解：(1) 检验统计量为：

$$U=\sqrt{n}\frac{\overline{X}-\mu_0}{\sigma}=\frac{\overline{X}-0.5}{0.1}$$

由题意，拒绝域为$\{U>u_\alpha=u_{0.05}=1.6449\}$

$$\Longrightarrow\quad\{\overline{X}>0.66449\}$$

(2) 直接列式：$\sigma_0=0.4$；$\mu=0.65$；$u_\alpha=u_{0.05}=1.6449$；$\mu_0=0.5$

$$\beta=P\left(\overline{X}<\mu_0+\frac{\sigma_0}{\sqrt{n}}u_\alpha\right)=P\left(\sqrt{n}\frac{\overline{X}-\mu}{\sigma_0}<\sqrt{n}\frac{\left(\mu_0+\frac{\sigma_0}{\sqrt{n}}u_\alpha\right)-\mu}{\sigma_0}\right)=\Phi\left(\sqrt{n}\frac{\left(\mu_0+\frac{\sigma_0}{\sqrt{n}}u_\alpha\right)-\mu}{\sigma_0}\right)$$

$$=\Phi(0.145)=0.560$$

犯第二类错误的概率约为 56.0%

8.4. 假设随机变量 X 的概率密度函数为 $f(x)=(1+\theta)x^\theta$，$0<x<1$，现考虑假设检验问题：

$H0$：$\theta=5$ $\leftrightarrow$ $H1$：$\theta=3$.

拒绝域为 $\{X>1/2\}$. 试求该检验问题的 I 类和 II 类错误，以及 $\theta=2$时的功效函数值.

解：功效函数 $\beta(\theta)=P_\theta(H_0\text{被拒绝})$

先写出$\theta=2$时功效函数的值：

$$\beta(\theta)=P_{\theta=2}\left(x>\frac{1}{2}\right)=\int_{\frac{1}{2}}^{1}3x^2\,dx=\frac{7}{8}$$

再写出犯 I 类和 II 类错误时功效函数的值：

第 I 类错误：$\beta(\theta)=P_{\theta\in H_0}(H_0\text{被拒绝})=P_{\theta=5}\left(x>\frac{1}{2}\right)=\int_{\frac{1}{2}}^{1}6x^5\,dx=\frac{63}{64}$

第 II 类错误：$1-\beta(\theta)=1-P_{\theta\in H_1}(H_0\text{被拒绝})=1-P_{\theta=3}\left(x>\frac{1}{2}\right)=1-\int_{\frac{1}{2}}^{1}4x^3\,dx=\frac{1}{16}$

8.5. 设样本 $X_1,\cdots,X_n$ 抽自参数为λ的泊松分布总体，对检验问题

$$H0:\quad\lambda=\frac{1}{2}\quad\leftrightarrow\quad H1:\quad\lambda\neq\frac{1}{2}$$

取检验的拒绝域为 $\{(X_1,\cdots,X_n):\sum_{i=1}^{10}X_i\leq 1\ or\geq 12\}$

(1) 求此检验在 $\lambda=0.25$，0.5，1 处的功效函数值，并求出该检验的水平.

(2) 求犯第一类错误的概率 及 在$\lambda=0.25$，0.75 处犯第二类错误的概率.

解：(1) 由题意可知：$X\sim P(\lambda)$，并令$Y=X_1+X_2+\cdots\cdots+X_n$，$Y\sim P(10\lambda)$

$$\Longrightarrow\quad P(Y=k)=\frac{(10\lambda)^k}{k!}e^{-10\lambda}\text{，}k\in\mathbb{N}$$

功效函数$\beta(\lambda)=P_\lambda(H_0\text{被拒绝})=P_\lambda(Y\leq 1\ or\ Y\geq 12)=1-P_\lambda(1<Y<12)=1-e^{-10\lambda}\sum_{i=2}^{11}\frac{(10\lambda)^i}{i!}$

再将三个待计算的λ值带入，即可得：

$$\lambda=0.25\text{：}\quad\beta(\lambda)=1-e^{-10*0.25}\sum_{i=2}^{11}\frac{(10*0.25)^i}{i!}=0.28731$$

$$\lambda=0.5\text{：}\quad\beta(\lambda)=1-e^{-10*0.5}\sum_{i=2}^{11}\frac{(10*0.5)^i}{i!}=0.04588$$

$$\lambda=1\text{：}\quad\beta(\lambda)=1-e^{-10*1}\sum_{i=2}^{11}\frac{(10*1)^i}{i!}=0.30372$$

(2) 犯第一类错误的概率：$\lambda=0.5$

$$P_{0.5}(H_0\text{被拒绝})=1-e^{-10*0.5}\sum_{i=2}^{11}\frac{(10*0.5)^i}{i!}=0.04588$$

再计算λ取不同值的时候犯第二类错误的概率：记犯第二类错误这个事件为事件A

$$\lambda=0.25\text{：}\quad P(A)=1-P_{0.25}(H_0\text{被拒绝})=1-0.28731=0.71269$$

$$\lambda=0.75\text{：}\quad P(A)=1-P_{0.75}(H_0\text{被拒绝})=e^{-10*0.75}\sum_{i=2}^{11}\frac{(10*0.75)^i}{i!}=0.91606$$