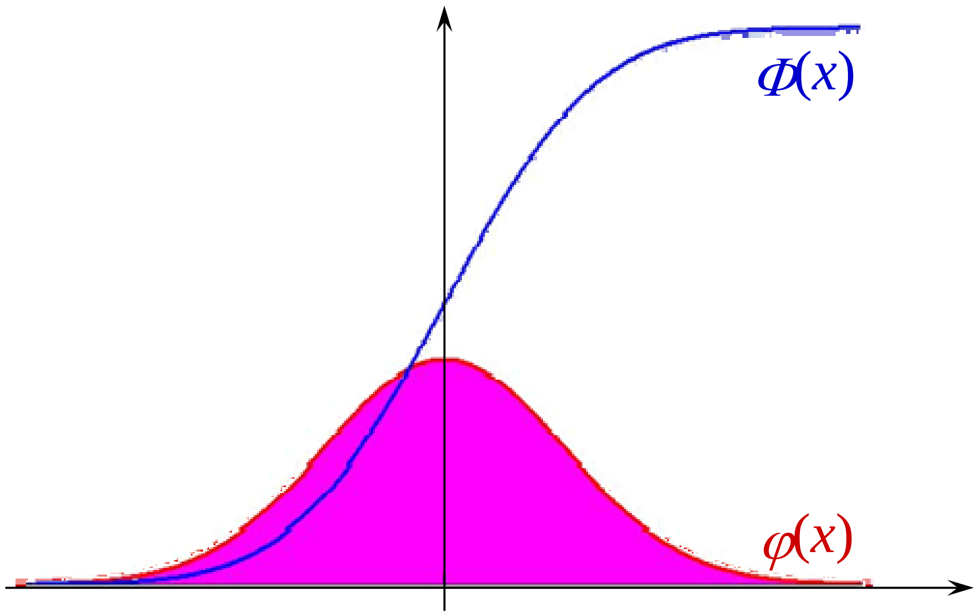
湖北省精品课程

概率论与数理统计



华中科技大学 概率统计系

叶 鹰 副教授

§7.4 区间估计

7.4.1 概念

问题：如何让$\hat{\theta}$与θ的误差体现在估计中？

办法：对给定的置信水平(置信度)$1-\alpha$

由样本$(X_1, X_2, \ldots, X_n)$构造：

置信下限 $\hat{\theta}_1(X_1,X_2,\ldots,X_n)$和置信上限 $\hat{\theta}_2(X_1,X_2,\ldots,X_n)$

使

$$P(\hat{\theta}_1 < \theta < \hat{\theta}_2) = 1-\alpha$$

称$(\hat{\theta}_1, \hat{\theta}_2)$为未知参数$\theta$的置信度为$1-\alpha$的置信区间。

含义：

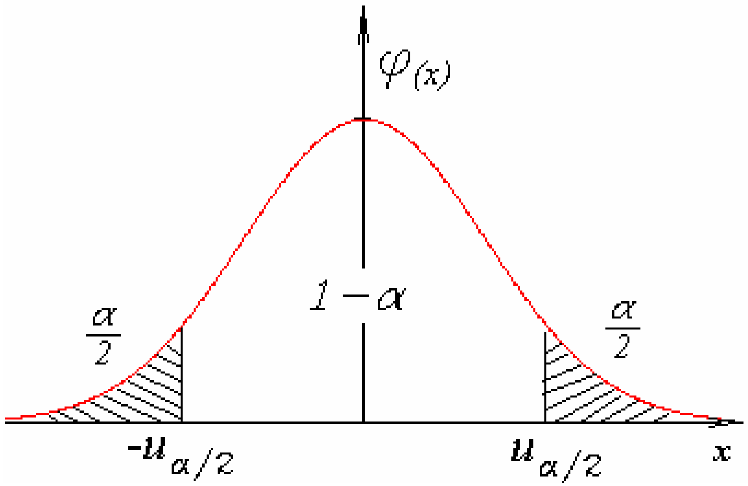
若$1-\alpha=0.95$，抽样100次中约有95个$(\hat{\theta}_1, \hat{\theta}_2)$包含$\theta$。

7.4.2 $N(\mu,\ \sigma^2)$中μ 的置信区间

1、 σ^2已知

$$\frac{\overline{X}-\mu}{\sqrt{\sigma^2/n}} \sim N(0,1)$$

$$P\left(\frac{\left|\overline{X}-\mu\right|}{\sigma/\sqrt{n}} < u_{\alpha/2}\right) = 1-\alpha$$



$$P\left(\overline{X}-\frac{\sigma}{\sqrt{n}}u_{\alpha/2} < \mu < \overline{X}+\frac{\sigma}{\sqrt{n}}u_{\alpha/2}\right) = 1-\alpha$$

即参数μ 的置信度为1－α的置信区间为

$$\left(\overline{X}-\frac{\sigma}{\sqrt{n}}u_{\alpha/2}\,,\ \overline{X}+\frac{\sigma}{\sqrt{n}}u_{\alpha/2}\right) \quad 或 \quad \left(\overline{X}\pm\frac{\sigma}{\sqrt{n}}u_{\alpha/2}\right)$$

例1（P_{128}例7.17）**滚珠直径**$X \sim N(\mu, 0.0006)$

$n = 6$ ：　　**1.46　1.51　1.49　1.48　1.52　1.51**

求μ的置信度为95%的置信区间。

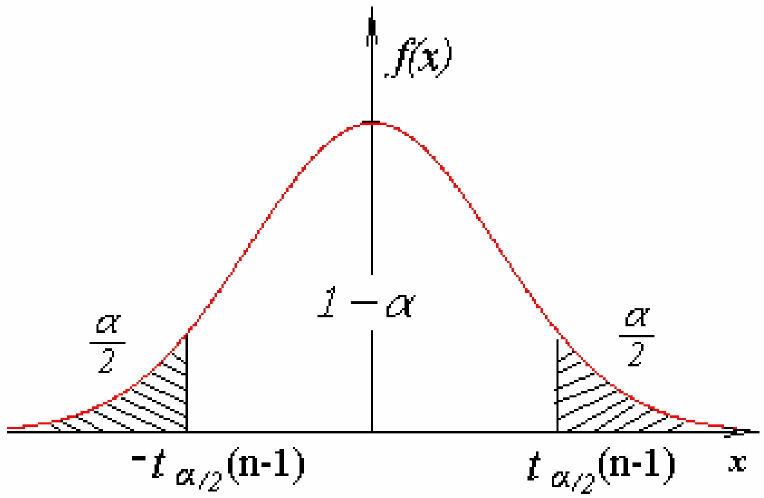
解　$\bar{x} = 1.495,\ \alpha = 0.05,\quad u_{0.05/2} = u_{0.025} = 1.96,$

$$\Rightarrow (1.495 \pm \sqrt{\frac{0.0006}{6}} \times 1.96) = (1.4754,\ 1.5146)$$

注：置信区间的长度： $l = 2 \times \frac{\sigma}{\sqrt{n}} u_{\alpha/2}$

（1）l与σ正比；　（2）l与$\sqrt{n}$反比；（3）α越大，l越小

2、σ^2未知



$$\frac{\bar{X}-\mu}{\sigma}\sqrt{n} \sim N(0,1)$$

$$T=\frac{\bar{X}-\mu}{S}\sqrt{n} \sim t(n-1)$$

$$P(|T|<t_{\alpha/2}(n-1))=1-\alpha$$

$$P(\bar{X}-\frac{S}{\sqrt{n}}t_{\alpha/2}(n-1)<\mu<\bar{X}+\frac{S}{\sqrt{n}}t_{\alpha/2}(n-1))=1-\alpha$$

即参数μ的置信度为$1-\alpha$的置信区间为 $(\bar{X}\pm\frac{S}{\sqrt{n}}t_{\alpha/2}(n-1))$

例2（续例1）若σ未知，则计算 s**=0.02258**，查表

$t_{0.025}(5)=2.5706$，算得

$\mu\in(1.495\pm0.0237)=(1.4716，1.5187)$

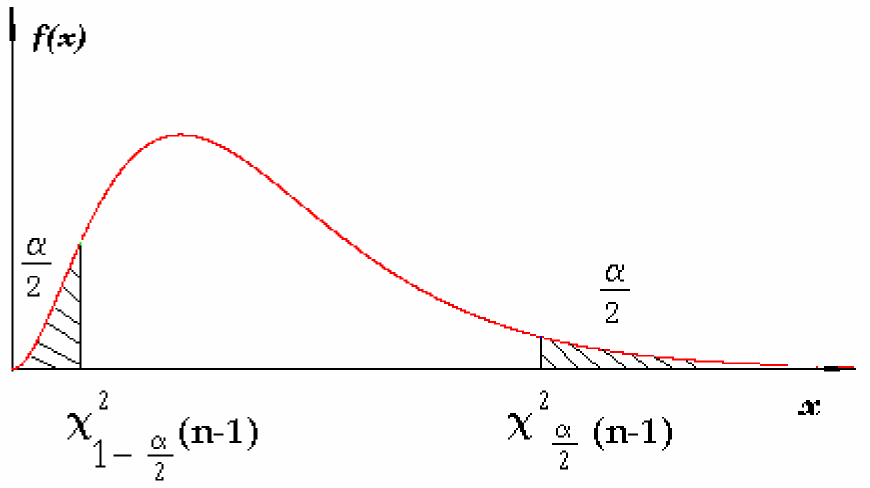
$(1.4754，1.5146)$

$\mathbf{l_2=0.0474}$

$\mathbf{>l_1=0.0392}$

7.4.3 $N(\mu，\sigma^2)$中σ^2的置信区间

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1) \qquad P\left(\chi^2_{1-\frac{\alpha}{2}}(n-1) < \frac{(n-1)S^2}{\sigma^2} < \chi^2_{\frac{\alpha}{2}}(n-1)\right) = 1-\alpha$$



$$P\left(\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\right) = 1-\alpha$$

例3（P_{131}例7.18）零件长度$X \sim N(\mu,\sigma^2)$，测得其中

n=16件： 2.15 2.10 2.12 2.10 2.14 2.11 2.15 2.13

2.13 2.11 2.14 2.13 2.12 2.13 2.10 2.14

求σ的95%置信区间。

解 计算（$\bar{x}=2.125$），$S^2=0.000293$，$\alpha=0.05$

查表 $\chi^2_{0.025}(15)=27.488$， $\chi^2_{0.975}(15)=6.262$

$$P\left(\sqrt{\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}}<\sqrt{\sigma^2}<\sqrt{\frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}}\right)=1-\alpha$$

$$\sqrt{\frac{15\times0.000293}{27.488}}<\sigma<\sqrt{\frac{15\times0.000293}{6.262}}$$

$$\Rightarrow \quad (0.01265, \quad 0.02651)$$

第八章 假设检验

§8.1 基本问题和方法

8.1.1 问题的提出

引例（P_{144}例8.4）某厂用一台包装机自动包装葡萄糖，包得的袋装糖重量服从 $N(\mu, \sigma^2)$。当机器正常时，其均值 μ_0= 0.5kg，标准差 $\sigma_0 = 0.015$kg，某日开工后随机地抽取9袋葡萄糖，称得重量（单位：kg）为0.497，0.506，0.518，0.524，0.498，0.511，0.520，0.515，0.512，问这台包装机是否正常?

- **总体：** 袋装味精重量 $X \sim N(\mu，\sigma^2)$
- **样本观察值：** **0.497，0.506，0.518，0.524，0.498，0.511，0.520，0.515，0.512**
- **正常：** $\mu = 0.5$ $\sigma = 0.015$
- **方法：** 假定正常 H_0：$\mu = 0.5$ 则

$$|\bar{x} - 0.5| < k$$

"太大"的标准 k =?

- **小概率原则：** 小概率事件在一次试验中不应该发生

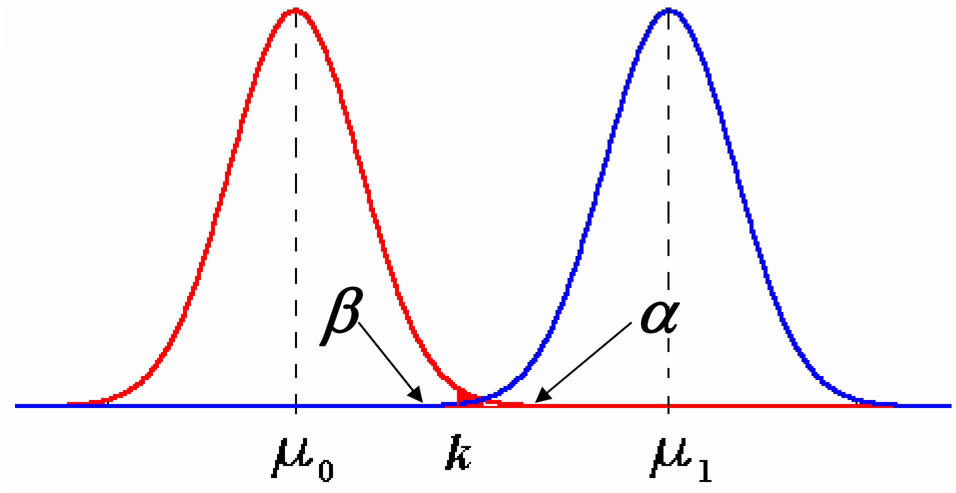
$$P(|\bar{X} - 0.5| > k) = \alpha \qquad （小概率）$$

- **检验统计量的分布：** $\bar{X} - 0.5 \sim N(0, \dfrac{\sigma^2}{n})$

8.1.2　两类错误

第一类错误 ~ 弃真：P（拒绝 | H_0为真）$=\alpha$

第二类错误 ~ 存伪：P（接受 | H_0不真）$=\beta$



办法：

•在控制$\alpha < \alpha_0$（显著水平）的前提下，使β尽可能地小。

•增加样本容量n，可同时减少α和β。

8.1.3 检验步骤

1、根据实际问题，提出原假设H_0和备择假设H_1，如

$$\mathbf{H}_0: \mu = \mu_0 \qquad \mathbf{H}_1: \mu \neq \mu_0$$

2、构造检验统计量 $\boldsymbol{T}(\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_n)$，使$H_0$为真时，$T$有确定的分布，如

$$T = \frac{\overline{X} - \mu_0}{S}\sqrt{n} \sim t(n-1)$$

3、对给定的显著水平α，确定H_0的拒绝域W，使

$P((X_1, X_2, \cdots, X_n) \in W) = \alpha$，如$W = \{(x_1, x_2, \cdots, x_n): T > t_\alpha(n-1)\}$

4、作出检验结论：

$(x_1, x_2, \cdots, x_n) \in W \Rightarrow$ 拒绝$\boldsymbol{H}_0$

否则 $\Rightarrow$ 不拒绝$\boldsymbol{H}_0$ ~认为$\boldsymbol{H}_0$与实际情况差异不显著。

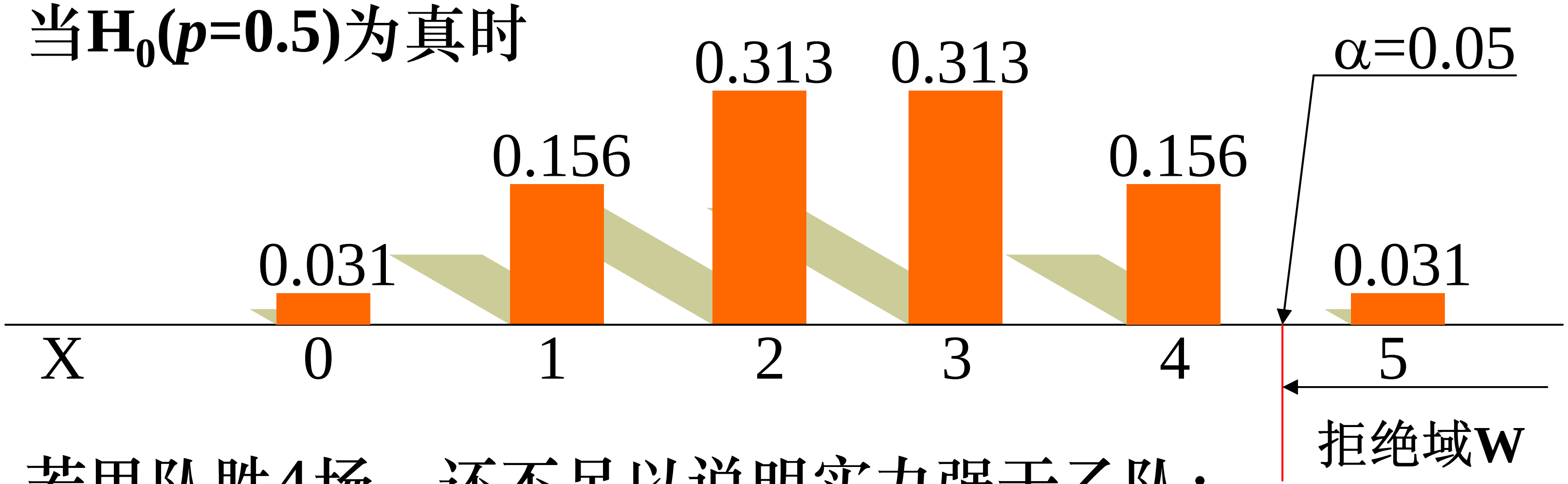
问题解答

在甲、乙两球队的5场比赛中，若甲队至少胜了4场，可否认为甲队实力明显强于乙队?

解 设X为甲队获胜的次数，则 $X \sim B(5,p)$.

$$H_0: p \leqslant 0.5,\quad H_1: p > 0.5$$

当$H_0(p=0.5)$为真时



若甲队胜4场，还不足以说明实力强于乙队；
若甲队胜5场，则可认为实力强于乙队。

习题选讲

练习12.6 将 r 个球随机放入 n 个盒子中，以X 记空盒子个数，求EX。

解 设n个随机变量如下：

$$X_i=\begin{cases}1, & \text{第}i\text{个盒子没球，}\\ 0, & \text{第}i\text{个盒子有球。}\end{cases} \qquad i=1,2,\cdots,n$$

则 $X=\sum_{i=1}^{n}X_i$ ，且

$$E(X_i)=P\{X_i=1\}=\frac{(n-1)^r}{n^r}$$

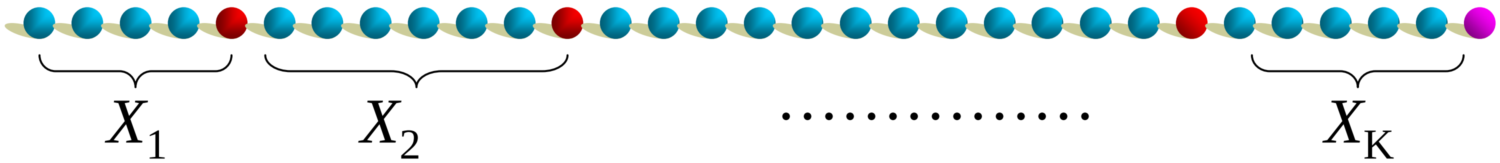
故

$$E(X)=\sum_{i=1}^{n}E(X_i)=n(1-\frac{1}{n})^r$$

习题选讲

练习13.4　流水作业线上生产的每一个产品不合格的概率为p，当生产出K个不合格产品时即停工检修一次，求两次检修之间产品总数的期望和方差。

解　设两次检修之间产品总数为X，考虑生产过程：



则$X_1, X_2, \ldots, X_K$独立同分布于几何分布。且

$$X=\sum_{i=1}^{K}X_i \qquad E(X_i)=\sum_{k=1}^{\infty}k(1-p)^{k-1}p=\frac{1}{p} \qquad D(X_i)=\frac{1-p}{p^2}$$

故　$$E(X)=\sum_{i=1}^{K}E(X_i)=\frac{K}{p} \qquad D(X)=\sum_{i=1}^{K}D(X_i)=K\frac{1-p}{p^2}$$

负二项分布：$P\{X=n\}=C_{n-1}^{k-1}p^k(1-p)^{n-k}$

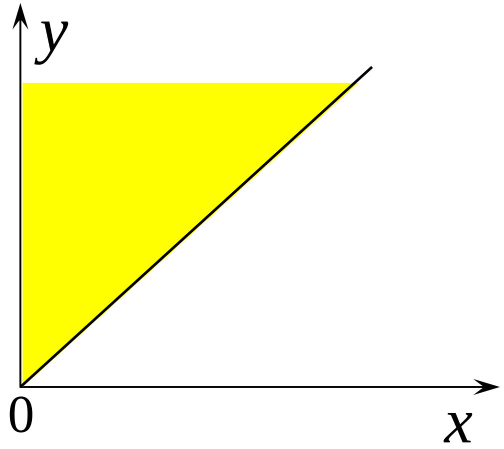
习题选讲

练习14.2 设 X 为取非负整数的离散型随机变量，其分布函数为 $F(x)$，试证明

$$EX=\boxed{\sum_{k=0}^{\infty}P\{X>k\}}$$

证明 设X的密度函数为 $f(x)$，则

$$\int_0^{+\infty}[1-F(x)]dx=\int_0^{+\infty}P\{X>x\}dx$$



$$=\int_0^{+\infty}[\int_x^{+\infty}f(y)dy]dx$$

$$=\int_0^{+\infty}[\int_0^{y}dx]f(y)dy$$

$$=\int_0^{+\infty}yf(y)dy=EX$$

习题选讲

习题4.7 设一个试验有m个等可能的结果，求至少一个结果连接发生k次的独立试验的期望次数。

解 设X_k为有一个结果接连发生k次的实验次数，

记 $E_k=E(X_k)$。则

$$X_k = X_{k-1} + [1\times\frac{1}{m} + (1-\frac{1}{m})X_k]$$

$$E_k = E_{k-1} + 1\times\frac{1}{m} + (1-\frac{1}{m})E_k \Rightarrow E_k = mE_{k-1} + 1$$

$$E_1 = 1 \Rightarrow E_k = 1 + m + m^2 + \cdots + m^{k-1} = \frac{m^k - 1}{m-1} \quad (m>1)$$