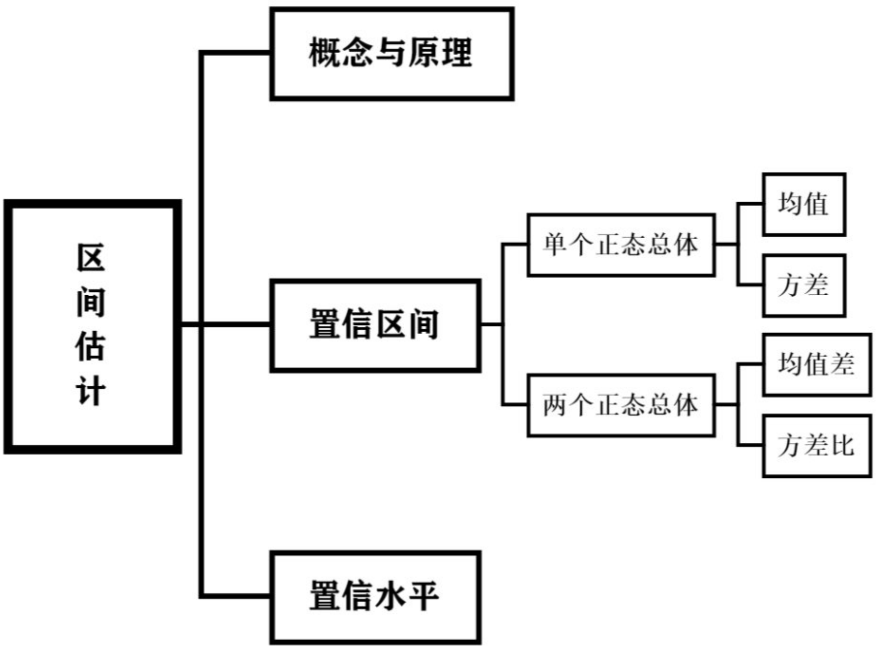
第八章 区间估计

知识图解



知识讲解

一、区间估计

1.定义： 设总体 X 的分布中含有一个未知参数 θ. 若对于给定的概率 $1-\alpha(0<\alpha<1)$，存在两个统计量 $\hat{\theta}_1=\hat{\theta}_1\left(X_1,X_2,\cdots,X_n\right)$ 与 $\hat{\theta}_2=\hat{\theta}_2\left(X_1,X_2,\cdots,X_n\right)$，使得

$$P\left\{\hat{\theta}_1\le\theta\le\hat{\theta}_2\right\}=1-\alpha$$

则随机区间 $\left(\hat{\theta}_1,\hat{\theta}_2\right)$ 称为参数 θ 的置信水平(或置信度)为 $1-\alpha$ 的置信区间（或区间估计），$\hat{\theta}_1$ 称为置信下限，$\hat{\theta}_2$ 称为置信上限，$1-\alpha$ 称为置信水平.

注：

1) 对于连续型总体可使用开区间，即$P\left\{\hat{\theta}_1<\theta<\hat{\theta}_2\right\}=1-\alpha$.

2) 置信水平$1-\alpha$越大，估计的可靠性越高.

3) 置信区间的含义：若反复抽样多次（各次的样本容量相等，均为n），每一组样本值确定一个区间$\left(\hat{\theta}_1,\hat{\theta}_2\right)$，每个这样的区间要么包含$\theta$的真值，要么不包含$\theta$的真值. 按伯努利大数定理，在这么多的区间中，包含θ真值的约占$100(1-\alpha)\%$.

4) 在一些实际问题中，人们只关心未知参数的一个下限或上限. 比如某种产品的寿命（希望越大越好），有$P\left\{\theta>\hat{\theta}\right\}=1-\alpha$，称$\hat{\theta}$为$\theta$置信水平为$1-\alpha$的单侧置信下限. 同理有单侧置信上限.

【例8.1】 设$X_1,X_2,\ldots,X_n$是来自正态总体$N\left(\mu,\sigma^2\right)$的样本，其中$\sigma^2$已知，设样本均值为$\bar{X}$，则$\mu$的置信度为$1-\alpha$的置信区间是________.（枢轴量法）

注：枢轴量$G\left(X_i,\mu\right)$为不含未知参数的关于样本和待估参数μ的函数.

【例8.2】 设$X_1,X_2,\ldots,X_n$是来自正态总体$N\left(\mu,\sigma^2\right)$的样本，其中$\sigma^2$未知，设样本均值为$\bar{X}$，样本方差为$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$，则$\mu$的置信度为$1-\alpha$的置信区间是________.

【例8.3】　设 $X_1, X_2, \ldots, X_n$ 是来自正态总体 $N\left(\mu, \sigma^2\right)$ 的样本，其中 μ 已知，则 σ^2 的置信度为 $1-\alpha$ 的置信区间是________.

【例8.4】　设 $X_1, X_2, \ldots, X_n$ 是来自正态总体 $N\left(\mu, \sigma^2\right)$ 的样本，其中 μ 已知，设样本均值为 $\bar{X}$，样本方差为 $S^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$，则 σ^2 的置信度为 $1-\alpha$ 的置信区间是________.

【例8.5】　设 $X_1, X_2, \ldots, X_m$ 是来自 $N\left(\mu_1, \sigma_1^2\right)$ 的样本，$Y_1, Y_2, \ldots, Y_n$ 是来自 $N\left(\mu_2, \sigma_2^2\right)$ 的样本，且两个样本相互独立. 设样本均值分别为 $\bar{X}$ 和 $\bar{Y}$，样本方差分别为 $S_X^2 = \dfrac{1}{m-1}\sum_{i=1}^{m}(X_i - \bar{X})^2$，$S_Y^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2$，则 σ_1^2 / σ_2^2 的置信度为 $1-\alpha$ 的置信区间是________.

2. 正态总体参数的双侧区间估计表（均取上分位点）

	待估参数	条件	统 计 量	双侧置信区间
一个总体	均值μ	σ^2已知	$\dfrac{\bar{X}-\mu}{\sigma_0 / \sqrt{n}} \sim N(0,1)$	$\left(\bar{X}-z_{\alpha/2}\cdot\dfrac{\sigma}{\sqrt{n}}, \bar{X}+z_{\alpha/2}\cdot\dfrac{\sigma}{\sqrt{n}}\right)$
		σ^2未知	$\dfrac{\bar{X}-\mu}{S / \sqrt{n}} \sim t(n-1)$	$\left(\bar{X}-t_{\alpha/2}(n-1)\cdot\dfrac{S}{\sqrt{n}}, \bar{X}+t_{\alpha/2}(n-1)\cdot\dfrac{S}{\sqrt{n}}\right)$
	方差σ^2	μ已知	$\dfrac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2 \sim \chi^2(n)$	$\left(\dfrac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{\alpha/2}(n)}, \dfrac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{1-\alpha/2}(n)}\right)$
		μ未知	$\dfrac{n-1}{\sigma_0^{\ 2}}S^2=\dfrac{1}{\sigma_0^{\ 2}}\sum_{i=1}^{n}(X_i-\bar{X})^2 \sim \chi^2(n-1)$	$\left(\dfrac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}, \dfrac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\right)$
两个总体	均值差 μ_1-μ_2	σ_1,σ_2均已知	$\dfrac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1}-\dfrac{\sigma_2^2}{n_2}}} \sim N(0,1)$	$\left((\bar{X}-\bar{Y})-\mu_{\alpha/2}\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2},\right.$ $\left.(\bar{X}-\bar{Y})+\mu_{\alpha/2}\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}\right)$
	均值差 μ_1-μ_2	σ_1,σ_2均未知 但σ_1=σ_2	$\dfrac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sqrt{\dfrac{n_1S_1^2+n_2S_2^2}{n_1+n_2-2}\cdot\dfrac{n_1+n_2}{n_1n_2}}} \sim t(n_1+n_2-2)$ $\dfrac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}} \sim t(n_1+n_2-2)$ $S_w^2=\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$	$\left((\bar{X}-\bar{Y})-t_{\alpha/2}(n_1+n_2-2)\cdot S_w\sqrt{1/n_1+1/n_2},\right.$ $\left.(\bar{X}-\bar{Y})+t_{\alpha/2}(n_1+n_2-2)\cdot S_w\sqrt{1/n_1+1/n_2}\right)$
	方差比 σ_1^2/σ_2^2	μ_1，μ_2 均未知	$\dfrac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F(n_1-1,n_2-1)$	$\left(\dfrac{S_1^2}{S_2^2}\cdot\dfrac{1}{F_{\alpha/2}(n_1-1,n_2-1)}, \dfrac{S_1^2}{S_2^2}\cdot\dfrac{1}{F_{1-\alpha/2}(n_1-1,n_2-1)}\right)$

【例8.6】 设一批零件的长度服从正态分布$N(\mu,\sigma^2)$，其中μ,σ^2均未知.现从中随机抽取16个零件，测得样本均值$\bar{x}=20(cm)$，样本标准差$s=1(cm)$，则μ的置信度为0.90的置信区间是（ ）

（A）$\left(20-\dfrac{1}{4}t_{0.05}(16), 20+\dfrac{1}{4}t_{0.05}(16)\right)$　　（B）$\left(20-\dfrac{1}{4}t_{0.1}(16), 20+\dfrac{1}{4}t_{0.1}(16)\right)$

（C）$\left(20-\dfrac{1}{4}t_{0.05}(15), 20+\dfrac{1}{4}t_{0.05}(15)\right)$　　（D）$\left(20-\dfrac{1}{4}t_{0.1}(15), 20+\dfrac{1}{4}t_{0.1}(15)\right)$

【例8.7】 设总体 $X \sim N\left(\mu,\sigma^2\right)$，来自 X 的一个简单随机样本 $X_1,X_2,\ldots,X_n$，μ 已知，未知参数 σ^2 的置信度 $1-\alpha$ 的置信区间为（ ）

（A）$\left[\dfrac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{1-\frac{\alpha}{2}}\left(n\right)},\dfrac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{\frac{\alpha}{2}}\left(n\right)}\right]$ （B）$\left[\dfrac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{\frac{\alpha}{2}}\left(n\right)},\dfrac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{1-\frac{\alpha}{2}}\left(n\right)}\right]$

（C）$\left[\dfrac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{1-\frac{\alpha}{2}}\left(n-1\right)},\dfrac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{\frac{\alpha}{2}}\left(n-1\right)}\right]$ （D）$\left[\dfrac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{\frac{\alpha}{2}}\left(n-1\right)},\dfrac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{1-\frac{\alpha}{2}}\left(n-1\right)}\right]$

【例8.8】 设轴承内环锻压零件的平均高度 X 服从正态分布 $N(\mu,0.4^2)$．现从中抽取 20 只内环，其平均高度 $\bar{x}=32.3\text{mm}$，求内环平均高度的置信度为 95% 的置信区间.

【例8.9】 设冷铜丝的折断力服从正态分布，从一批铜丝中任取10根，测试折断力，得数据为 578，572，570，568，572，570，570，596，584，572

求：(1) 样本均值和样本方差；(2) 方差的置信区间（$\alpha=0.05$）.

【例8.10】 设总体 $X \sim N\left(\mu,\sigma^2\right)$，来自 X 的一个简单随机样本 $X_1,X_2,\dots,X_n$，记 $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$，$T=\sum_{i=1}^{n}(X_i-\overline{X})^2+\sum_{i=n+1}^{2n}(X_i-\mu)^2$，当 μ 已知时，基于 T 构造估计 σ^2 的置信水平为 $1-\alpha$ 的置信区间为（ ）

（A）$\left(\frac{T}{\chi^2_{1-\frac{\alpha}{2}}\left(2n\right)},\frac{T}{\chi^2_{\frac{\alpha}{2}}\left(2n\right)}\right)$ （B）$\left(\frac{T}{\chi^2_{\frac{\alpha}{2}}\left(2n\right)},\frac{T}{\chi^2_{1-\frac{\alpha}{2}}\left(2n\right)}\right)$

（C）$\left(\frac{T}{\chi^2_{1-\frac{\alpha}{2}}\left(2n-1\right)},\frac{T}{\chi^2_{\frac{\alpha}{2}}\left(2n-1\right)}\right)$ （D）$\left(\frac{T}{\chi^2_{\frac{\alpha}{2}}\left(2n-1\right)},\frac{T}{\chi^2_{1-\frac{\alpha}{2}}\left(2n-1\right)}\right)$

【例8.11】 某大学从来自 A，B 两市的新生中分别随机抽取 5 名与 6 名新生，测其身高（单位：cm）后算得 $\overline{x}=175.9, \overline{y}=172.0, s_1^2=11.3, s_2^2=9.1$. 假设两市新生身高分别服从正态分布 $X \sim N\left(\mu_1,\sigma^2\right), Y \sim N\left(\mu_2,\sigma^2\right)$，其中 σ^2 未知. 试求 $\mu_1-\mu_2$ 的置信度为 0.95 的置信区间. （$t_{0.025}(9)=2.2622, t_{0.025}(11)=2.2010$）

【例8.12】 某车间有两条流水线加工一类表盘，假设表盘直径服从正态分布. 现从两条流水线分别抽查 5 个和 6 个表盘，得直径（单位：cm）数据如下：

1 号流水线：5.06，5.08，5.03，5.00，5.07

2 号流水线：4.98，5.03，4.97，4.99，5.02，4.95

试求两条流水线加工表盘直径的方差比 σ_1^2/σ_2^2 的置信度为 0.95 的置信区间.

（$F_{0.975}(4,5)=\dfrac{1}{F_{0.025}(5,4)}=\dfrac{1}{9.36}$，$F_{0.025}(4,5)=7.39$）

【例8.13】 在测量反应时间（服从正态分布）的试验中，一心理学家估计的标准差为$0.05s$，为了以95%的置信度使他对平均反应时间的估计误差不超过$0.01s$，应取多大的样本容量n？