§5 正态总体均值与方差的区间估计

- 一个正态总体的未知参数的置信区间
- 两个正态总体中未知参数的置信区间

求置信区间的步骤：

(1) 找一个枢轴量 $W=W(X_1,\cdots,X_n;\theta)$.

(2) 对给定的置信水平 $1-\alpha$，确定常数 a,b, 使

$$P\{a<W(X_1,\cdots,X_n;\theta)<b\}=1-\alpha.$$

(3) $a<W(X_1,\cdots,X_n;\theta)<b \Leftrightarrow \quad \underline{\theta}<\theta<\overline{\theta}$,

其中 $\underline{\theta}=\underline{\theta}(X_1,\cdots,X_n)$，$\overline{\theta}=\overline{\theta}(X_1,\cdots,X_n)$ 都是统计量 .

(4) 随机区间 $(\underline{\theta},\ \overline{\theta})$ 是 θ 的一个置信水平为 $1-\alpha$ 的置信区间.

一、一个正态总体的未知参数的置信区间

1） 均值的区间估计

设$X_1,\cdots,X_n$为总体 $X\sim N(\mu,\sigma^2)$ 的一个样本,

在置信水平$1-\alpha$下，来确定μ的置信区间$(\underline{\theta},\ \overline{\theta})$.

(1)方差已知时，估计均值

设已知方差$\sigma^2=\sigma_0^2$,

构造样本的函数 $Z=\dfrac{\overline{X}-\mu}{\sigma_0/\sqrt{n}}\sim N(0,1).$

μ的一个置信水平为$1-\alpha$的置信区间为

$$\left(\ \overline{X}\pm\frac{\sigma_0}{\sqrt{n}}z_{\alpha/2}\ \right)$$

(2) 方差未知时，估计均值

S^2是σ^2的无偏估计，

从而选取样本函数： $T=\dfrac{\overline{X}-\mu}{S/\sqrt{n}}\sim t(n-1).$

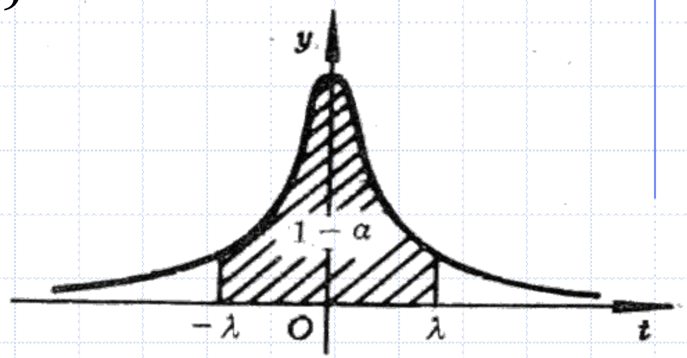
对于给定的 $1-\alpha$，查 t 分布表，找 λ_1 与 λ_2，使得：

$$P\{\lambda_1<T<\lambda_2\}=1-\alpha,$$

我们仍然取成对称区间 $(-\lambda,\lambda)$，使得：

$$P\{|T|<\lambda\}=1-\alpha,$$

由t分布表的构造及 $P\{|T|<\lambda\}=1-\alpha$，可知：



$\lambda = t_{\frac{\alpha}{2}}(n-1)$， 由此得：

$$-t_{\frac{\alpha}{2}}(n-1) < \frac{\overline{X}-\mu}{S/\sqrt{n}} < t_{\frac{\alpha}{2}}(n-1)$$

推得，置信区间为：

$$\left(\overline{X}-t_{\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}},\ \overline{X}+t_{\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}}\right)$$

例1 一家粮油公司大米包装生产线封装的袋米重量服从 $N(\mu,\sigma^2)$.从生产线抽取10袋大米,测得各袋大米的重量(单位:千克)如下:

10.1, 10.0, 9.8, 10.5, 9.7, 10.1, 9.9, 10.2, 10.3, 9.9.

求该生产线封装的平均袋米重量μ的置信水平为95%的置信区间。

解： 所用枢轴量为：
$$t=\frac{\bar{X}-\mu}{S/\sqrt{n}}\sim t(n-1).$$

μ的置信水平为$1-\alpha$的置信区间为

$$\left(\bar{X}-t_{\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}},\ \bar{X}+t_{\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}}\right)$$

已知 $n=10,\alpha=0.05$． 由样本值算得：

$$\bar{x}=10.05,\quad s^2=0.0583.$$

查表得 $t_{0.025}(9)=2.2622$，由此得 μ的95% 置信区间：

$$\left(10.05-2.2622\sqrt{\frac{0.0583}{9}},\ 10.05+2.2622\sqrt{\frac{0.0583}{9}}\right)$$

$$=(9.87,\ 10.22)$$

2）方差的区间估计

设$X_1,\cdots,X_n$为总体$X \sim N(\mu,\sigma^2)$的一个样本.

根据实际问题的需要,只介绍μ未知的情况.

枢轴量: $\chi^2 = \dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$.

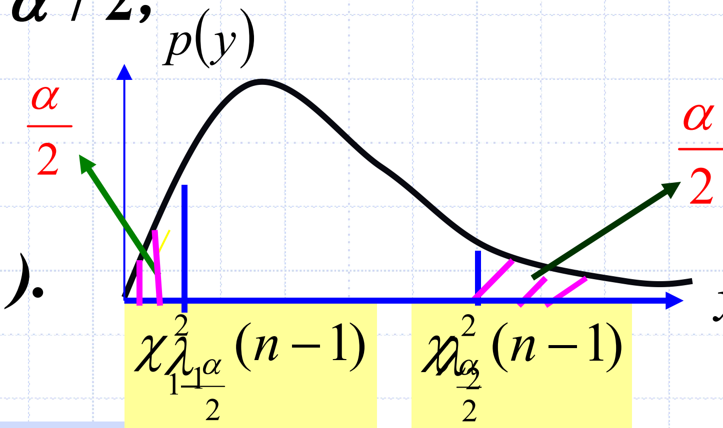
对于给定的 $1-\alpha$，查 χ^2分布表，得 λ_1与λ_2使得:

$$P\{\lambda_1 < \chi^2 < \lambda_2\} = 1-\alpha,$$

虽然χ^2分布密度函数无对称性，我们仍采用使概率对称的区间:

$$P\{\chi^2 < \lambda_1\} = P\{\chi^2 > \lambda_2\} = \alpha/2,$$

$$P\{\chi^2 < \lambda_1\} = P\{\chi^2 > \lambda_2\} = \alpha/2,$$



查$\chi^2(n-1)$分布表，得

$$\lambda_2 = \chi^2_{\frac{\alpha}{2}}(n-1), \quad \lambda_1 = \chi^2_{1-\frac{\alpha}{2}}(n-1).$$

由此得：

$$\chi^2_{1-\frac{\alpha}{2}}(n-1) < \frac{(n-1)S^2}{\sigma^2} < \chi^2_{\frac{\alpha}{2}}(n-1)$$

推得：

$$\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}$$

得方差σ^2的一个置信水平为$1-\alpha$的置信区间

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

标准差σ的一个置信水平为$1-\alpha$的置信区间

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

例2 设某机床加工的零件长度 $X \sim N(\mu,\sigma^2)$，今抽查16个零件，测得长度（单位：mm）如下：

12.15，12.12，12.01，12.08，12.09，12.16，12.03，12.01，
12.06，12.13，12.07，12.11，12.08，12.01，12.03，12.06，

在置信水平为95%时，试求总体方差σ^2的置信区间.

解： 已知$n=16, \alpha=0.05$.由样本值算得: $s^2=0.00244$.

查$\chi^2(n-1)$分布表,得 $\chi^2_{0.025}(15)=27.5$,

$$\chi^2_{0.975}(15)=6.26.$$

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

由此得置信区间:

$$\left(\frac{15\times 0.00244}{27.5}, \frac{15\times 0.00244}{6.26}\right)=(0.0013,\ 0.0058)$$

一个正态总体未知参数的置信区间

待估参数		枢轴量	枢轴量的分布	双侧置信区间的上、下限
μ	σ^2已知	$\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}$	$N(0,\ 1)$	$\overline{X}\pm z_{\frac{\alpha}{2}}\cdot\dfrac{\sigma}{\sqrt{n}}$
	σ^2未知	$\dfrac{\overline{X}-\mu}{S/\sqrt{n}}$	$t(n-1)$	$\overline{X}\pm t_{\frac{\alpha}{2}}(n-1)\cdot\dfrac{S}{\sqrt{n}}$
σ^2	μ已知 (*)	$\dfrac{1}{\sigma^2}\sum\limits_{i=1}^{n}(X_i-\mu)^2$	$\chi^2(n)$	$\dfrac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{\frac{\alpha}{2}}(n)}$ $\dfrac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{1-\frac{\alpha}{2}}(n)}$
	μ未知	$\dfrac{1}{\sigma^2}\sum\limits_{i=1}^{n}(X_i-\overline{X})^2=\dfrac{(n-1)S^2}{\sigma^2}$	$\chi^2(n-1)$	$\dfrac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}$ $\dfrac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}$

二、两个正态总体中未知参数的置信区间

设 $X_1,\cdots,X_{n_1}$ 是从正态总体 $X\sim N\left(\mu_1,\ \sigma_1^2\right)$ 中抽取的样本，

$Y_1,\cdots,Y_{n_2}$ 是从正态总体 $Y\sim N\left(\mu_2,\ \sigma_2^2\right)$ 中抽取的样本，

下表给出了置信水平为 $1-\alpha$ 的各种置信区间：

两个正态总体均值差 $\mu_1-\mu_2$ 的置信区间（一）

待估参数		枢轴量	枢轴量的分布	双侧置信区间的上、下限
$\mu_1-\mu_2$	σ_1^2、σ_2^2 均已知	$\dfrac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}$	$N(0,\ 1)$	$(\bar{X}-\bar{Y})\pm z_{\frac{\alpha}{2}}\cdot\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}$
	$\sigma_1^2=\sigma_2^2=\sigma^2$ 但未知	$\dfrac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}$	$t(n_1+n_2-2)$	$(\bar{X}-\bar{Y})\pm t_{\frac{\alpha}{2}}(n_1+n_2-2)S_w\cdot\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}$

其中 $S_w^2=\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$

两个正态总体方差比 $\dfrac{\sigma_1^2}{\sigma_2^2}$ 的置信区间（二）

待估参数		枢轴量	枢轴量的分布	双侧置信区间的上、下限
$\dfrac{\sigma_1^2}{\sigma_2^2}$	μ_1、μ_2 均已知 (*)	$\dfrac{n_2\sum_{i=1}^{n_1}(X_i-\mu_1)^2/\sigma_1^2}{n_1\sum_{j=1}^{n_2}(Y_j-\mu_2)^2/\sigma_2^2}$	$F(n_1,\ n_2)$	$\dfrac{1}{F_{\frac{\alpha}{2}}(n_1,\ n_2)}\cdot\dfrac{n_2\sum_{i=1}^{n_1}(X_i-\mu_1)^2}{n_1\sum_{j=1}^{n_2}(Y_j-\mu_2)^2}$, $\dfrac{1}{F_{1-\frac{\alpha}{2}}(n_1,\ n_2)}\cdot\dfrac{n_2\sum_{i=1}^{n_1}(X_i-\mu_1)^2}{n_1\sum_{j=1}^{n_2}(Y_j-\mu_2)^2}$
	μ_1、μ_2 均未知	$\dfrac{S_1^2\big/\sigma_1^2}{S_2^2\big/\sigma_2^2}$	$F(n_1-1,\ n_2-1)$	$\dfrac{1}{F_{\frac{\alpha}{2}}(n_1-1,\ n_2-1)}\cdot\dfrac{S_1^2}{S_2^2}$, $\dfrac{1}{F_{1-\frac{\alpha}{2}}(n_1-1,\ n_2-1)}\cdot\dfrac{S_1^2}{S_2^2}$

例3 为比较甲乙两类试验田的收获量，随机抽取甲类试验田 8 块，乙类试验田 10 块，分别测得其收获量如下（单位：kg）：

甲类：12.6， 10.2， 11.7， 12.3， 11.1， 10.5， 10.6， 12.2，

乙类：8.6， 7.9， 9.3， 10.7， 11.2， 11.4， 9.8， 9.5， 10.1， 8.5，

假设两类试验田的收获量都服从正态分布，且方差相同．试在置信水平 $1-\alpha=0.95$ 下，求两个总体均值差 $\mu_1-\mu_2$ 的置信区间．

解： 由上表，取枢轴量

$$T=\frac{\left(\overline{X}-\overline{Y}\right)-\left(\mu_1-\mu_2\right)}{\sqrt{\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\sim t(n_1+n_2-2)$$

$$\left(\overline{X}-\overline{Y}\right)\pm t_{\frac{\alpha}{2}}(n_1+n_2-2)S_w\cdot\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

$$\left(\overline{X}-\overline{Y}\right)\pm t_{\frac{\alpha}{2}}\left(n_1+n_2-2\right)S_w\cdot\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

由样本观测值

$n_1=8$， $\bar{x}=11.4$， $s_1^2=0.851$，

$n_2=10$， $\bar{y}=9.7$， $s_2^2=1.387$，

又由$1-\alpha=0.95$，得 $\frac{\alpha}{2}=0.025$

查表，得 $t_{\frac{\alpha}{2}}\left(n_1+n_2-2\right)=t_{0.025}(16)=2.12$

将上面各数代入置信区间端点的计算公式，得

$$(\bar{x}-\bar{y})-t_{\frac{\alpha}{2}}(n_1+n_2-2)s_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}=(11.4-9.7)-2.12\times0.508\times\sqrt{\frac{1}{8}+\frac{1}{10}}=0.6$$

$$(\bar{x}-\bar{y})+t_{\frac{\alpha}{2}}(n_1+n_2-2)s_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}=(11.4-9.7)+2.12\times0.508\times\sqrt{\frac{1}{8}+\frac{1}{10}}=2.8$$

其中 $S_w^2=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$

所求置信区间为(0.6, 2.8).

例4 研究由机器 A 与机器 B 生产的钢管的内径，随机抽取机器 A 生产的产品 18 只，测得其样本方差为 $s_1^2=0.34$（mm^2）；随机抽取机器 B 生产的产品 13 只，测得其样本方差为 $s_2^2=0.29$（mm^2）．假设两样本相互独立，且设由两机器生产的钢 管的内径分别服从正态分布$N(\mu_1,\sigma_1^2)$，$N(\mu_2,\sigma_2^2)$,其中$\mu_i,\sigma_i^2(i=1,2)$均未知. 试求方差比 σ_1^2/σ_2^2 的置信水平为0.90的置信区间.

解： 由题意，取枢轴量

$$F=\frac{S_1^2\Big/\sigma_1^2}{S_2^2\Big/\sigma_2^2}\sim F(n_1-1,n_2-1),$$

由
$$P\left\{ a < \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} < b \right\} = 1-\alpha$$

得 $a = F_{1-\frac{\alpha}{2}}(n_1-1,\ \ n_2-1)$，$b = F_{\frac{\alpha}{2}}(n_1-1,\ \ n_2-1)$，

因此，得 $\frac{\sigma_1^2}{\sigma_2^2}$ 的双侧 $1-\alpha$ 置信区间为

$$\left(\frac{1}{F_{\frac{\alpha}{2}}(n_1-1, n_2-1)} \cdot \frac{S_1^2}{S_2^2},\ \frac{1}{F_{1-\frac{\alpha}{2}}(n_1-1, n_2-1)} \cdot \frac{S_1^2}{S_2^2} \right)$$

由样本观测值 $n_1 = 18$， $s_1^2 = 0.34$，

$n_2 = 13$， $s_2^2 = 0.29$，

又由 $\mathbf{1-\alpha = 0.90}$，得 $\mathbf{1-\frac{\alpha}{2} = 0.95}$， $\mathbf{\frac{\alpha}{2} = 0.05}$

查表，得 $F_{\frac{\alpha}{2}}(n_1-1,\ \ n_2-1) = F_{0.05}(17,\ \ 12) = 2.59$

$$F_{1-\frac{\alpha}{2}}(n_1-1,\ \ n_2-1) = F_{0.95}(17,\ \ 12) = \frac{1}{F_{0.05}(12,\ \ 17)} = \frac{1}{2.38}$$

将上面各数代入置信区间端点的计算公式，得

$$\frac{1}{F_{\frac{\alpha}{2}}(n_1-1,\ \ n_2-1)}\cdot\frac{S_1^2}{S_2^2} = 0.45,\qquad \frac{1}{F_{1-\frac{\alpha}{2}}(n_1-1,\ \ n_2-1)}\cdot\frac{S_1^2}{S_2^2} = 2.79,$$

所求置信区间为 $(0.45,\ \ 2.79)$.

主要内容:

掌握置信区间的求法—— 选定枢轴量,

利用枢轴量的分布的上α 分位点。

§7 单侧置信区间

在某些实际问题中,例:对设备、元件的寿命来说，平均寿命长是我们所期望的，只关心平均寿命的下限;

与之相反，考虑化学药品中杂质含量的均值μ 时，我们常关心参数μ的上限。

——引出单侧置信区间的概念。

1 单侧置信区间

定义(1) 对给定值 $\alpha(0<\alpha<1)$, 若由来自总体 X 的样本 $X_1,\cdots,X_n$ 确定的统计量 $\underline{\theta}=\underline{\theta}(X_1,\cdots,X_n), \forall\theta\in\Theta$, 满足

$$P\{\theta>\underline{\theta}(X_1,\cdots,X_n)\}\geq 1-\alpha$$

称区间 $(\underline{\theta},+\infty)$ 是 θ 的置信水平为 $1-\alpha$ 的单侧置信区间.

$\underline{\theta}$: 为 θ 的置信水平是 $1-\alpha$ 的单侧置信下限 .

(2)又若 $\overline{\theta}=\overline{\theta}(X_1,\cdots,X_n),\forall\theta\in\Theta$,有

$$P\{\theta<\overline{\theta}(X_1,\cdots,X_n)\}\geq 1-\alpha$$

称区间 $(-\infty,\overline{\theta})$ 是 θ 的置信水平为 $1-\alpha$ 的单侧置信区间 .

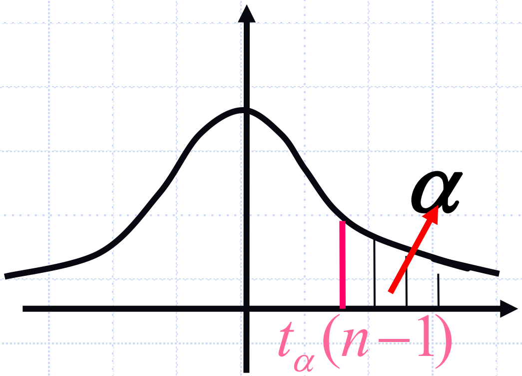
$\overline{\theta}$: 为 θ 的置信水平是 $1-\alpha$ 的单侧置信上限.

2 一个正态总体的未知参数的单侧置信区间

设$X_1,\cdots,X_n$为总体 $X\sim N(\mu,\sigma^2)$的一个样本，μ,σ^2均未知.

(1) μ的置信下限

$$\frac{\overline{X}-\mu}{S/\sqrt{n}}\sim t(n-1)$$



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

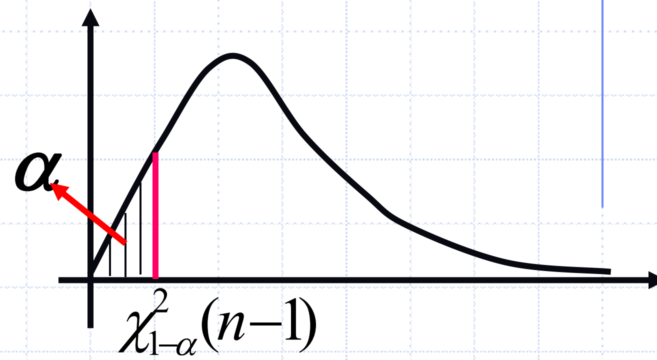
得μ的一个置信水平为$1-\alpha$的单侧置信区间

$$(\overline{X}-\frac{S}{\sqrt{n}}t_\alpha(n-1),+\infty)$$

置信下限

(2) σ^2的置信上限

由 $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$



$$P\{\frac{(n-1)S^2}{\sigma^2} > \chi^2_{1-\alpha}(n-1)\} = 1-\alpha$$

即： $P\{\sigma^2 < \dfrac{(n-1)S^2}{\chi^2_{1-\alpha}(n-1)}\} = 1-\alpha$

得σ^2的一个置信水平为$1-\alpha$的单侧置信区间

$$(0, \frac{(n-1)S^2}{\chi^2_{1-\alpha}(n-1)})$$

置信上限

例 从一批灯泡中随机地取5只作寿命试验，测得寿命(以小时计)为：**1050, 1100, 1120, 1250, 1280**；设灯泡寿命服从正态分布。求灯泡寿命平均值的置信水平为**0.95**的单侧置信下限。

解：因μ的一个置信水平为0.95的单侧置信区间

$$(\overline{X}-\frac{S}{\sqrt{n}}t_{\alpha}(n-1),+\infty)$$

而 $1-\alpha=0.95, n=5, t_{\alpha}(n-1)=t_{0.05}(4)=2.1318$,

$$\overline{x}=1160, s^{2}=9950,$$

得所求单侧置信下限为 $\underline{\mu}=\overline{x}-\frac{s}{\sqrt{n}}t_{\alpha}(n-1)=1065.$

第七章 小　结

1 理解参数的点估计、估计量与估计值的概念。

2 掌握矩估计法和最大似然估计法。

3 了解估计量的无偏性、有效性和相合性的概念，并会验证估计量的无偏性、有效性。

4 理解区间估计的概念，会求单个正态总体的均值和方差的置信区间。