

概 率 论 与 数 理 统 计

Probability and Statistics

南 京 大 学 张 绍 群 & 徐 科

更新: December 8, 2025

Ch10: 参数估计

Interval Estimation

December 8, 2025

区间估计

点估计用样本统计统计量直接作为总体参数的估计值, 简洁却存在以偏概全的局限性. 但是 **推断一个参数落在某个区间比推断这个参数具体的数值要简单**. 因此, 我们可以推断总体的参数的是在某个区间范围内的. 比如通过 1 万人的样本预估出中国总体人口的平均身高是 165-175cm, 这就是区间估计.

区间的范围很大, 你可以预估身高是 165-175cm 之间, 也可以预估是 160-180cm 之间, 也可以是其他. 但你会看到, 前者相比后者预测准确的概率更低, 因为其预测的区间范围太窄, 而后者预测的区间范围更宽. 所以, 在进行区间估计的时候, 你会发现每一个预估的区间都对应一个预估的准确度. 前者被称为置信区间, 后者被称为置信度.

提纲：区间估计

- 置信区间与置信度
- 枢轴 (Pivot) 变量法
 - 单个正态总体参数： σ 已知时， μ 的置信区间
 - 单个正态总体参数： σ 未知时， μ 的置信区间
 - 单个正态总体参数： σ^2 的置信区间
 - 两个正态总体参数
 - 非正态分布的区间估计
- 单侧置信区间
- 区间估计与机器学习

置信区间与置信度

定义 0.1 设 X_1, X_2, \dots, X_n 是来自总体 X 的样本, 总体 X 的分布函数含未知参数 θ , 求统计量 $\hat{\theta}_1 = \hat{\theta}_1(X_1, X_2, \dots, X_n)$ 和 $\hat{\theta}_2 = \hat{\theta}_2(X_1, X_2, \dots, X_n)$ 使得

$$\Pr[\hat{\theta}_1 < \theta < \hat{\theta}_2] \geq 1 - \alpha$$

成立, 则称 $1 - \alpha$ 为置信度, $[\hat{\theta}_1, \hat{\theta}_2]$ 为 θ 的置信度为 $1 - \alpha$ 的置信区间.

Remarks:

- 置信区间 $\hat{\theta}_2 - \hat{\theta}_1$ 反映了估计精度: 长度越小, 精度越大.
- 置信度 α 反映了估计的可靠度: α 越小, 可靠度越高.
- 给定置信度 α , 置信区间 $[\hat{\theta}_1, \hat{\theta}_2]$ 的选取并不唯一确定. 通常选长度最小的一个区间.

枢轴变量法

构造未知参数 θ 的置信区间最常用的方法是枢轴变量法, 其步骤可以概括如下:

- 设法构造一个样本统计量 $W = W(X_1, X_2, \dots, X_n; \theta)$, 使得 W 的分布包含待估参数 θ , 但不依赖其它参数, 且函数 W 的分布已知, **W 称为枢轴变量.**
- 给定置信度 $1 - \alpha$, 根据 W 的分布找出临界值 a 和 b , 使得下式成立

$$\Pr[a < W < b] = 1 - \alpha .$$

- 根据 $a < W < b$, 解出 $\hat{\theta}_1 < \theta < \hat{\theta}_2$, 则 $[\hat{\theta}_1, \hat{\theta}_2]$ 为 θ 的置信度为 $1 - \alpha$ 的置信区间.

单个正态总体参数—— σ 已知时 μ 的置信区间

按照枢轴变量法, 有:

- 由于 μ 的点估计为 \bar{X} , 且 $\bar{X} \sim \mathcal{N}(\mu, \sigma^2/n)$, 因此枢轴变量可选为 $W = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1)$.
- 给定置信度 $1 - \alpha$, 根据 W 的分布找出临界值 a 和 b , 使得

$$\Pr[a < W < b] = 1 - \alpha .$$

- 根据正态分布的性质、对称性和上分位点可知

$$\Pr[W \geq \mu_{1-\alpha/2}] = 1 - \alpha/2 \quad \text{和} \quad \Pr[W \leq \mu_{\alpha/2}] = 1 - \alpha/2 .$$

求解可得 $a = \mu_{1-\alpha/2} = -\mu_{\alpha/2}$ 和 $b = \mu_{\alpha/2}$. 又因为 $W = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$, 于是有

$$\Pr[-\mu_{\alpha/2} < W < \mu_{\alpha/2}] = \Pr \left[\bar{X} - \frac{\sigma}{\sqrt{n}} \mu_{\alpha/2} < \mu < \bar{X} + \frac{\sigma}{\sqrt{n}} \mu_{\alpha/2} \right] = 1 - \alpha .$$

区间估计：例 0.1

例 0.1 用天平称量某物体的质量 9 次, 得平均值 $\bar{x} = 15.4(\text{g})$, 已知天平称量结果为正态分布, 其标准差为 $0.1(\text{g})$. 试求该物体质量的 0.95 置信区间.

解答：例 0.1

题目：用天平称量某物体的质量 9 次，得平均值 $\bar{x} = 15.4(\text{g})$ ，已知天平称量结果为正态分布，其标准差为 $0.1(\text{g})$ 。试求该物体质量的 0.95 置信区间。

解答：

- 根据题意知 $1 - \alpha = 0.95$, $\alpha = 0.05$, 查标准正态分布函数表知 $\mu_{\alpha/2} = \mu_{0.025} = 1.96$.
- 进一步有该物体质量的 0.95 置信区间为

$$\bar{X} \pm \frac{\sigma}{\sqrt{n}} \mu_{\alpha/2} = 15.4 \pm 1.96 \times 0.1 / \sqrt{9} = 15.4 \pm 0.0653.$$

从而该物体质量的 0.95 置信区间为 $[15.3347, 15.4653]$ 。

区间估计：例 0.2

例 0.2 设总体为正态分布 $\mathcal{N}(\mu, 1)$, 为得到 μ 的置信度为 0.95 的置信区间且区间长度不超过 1.2, 求样本容量应为多大?

解答：例 0.2

题目：设总体为正态分布 $\mathcal{N}(\mu, 1)$, 为得到 μ 的置信度为 0.95 的置信区间且区间长度不超过 1.2, 求样本容量应为多大?

解答:

- 根据题意知 $1 - \alpha = 0.95$, $\alpha = 0.05$, 查标准正态分布函数表知 $\mu_{\alpha/2} = \mu_{0.975} = 1.96$.
- 又因为 μ 的置信度为 0.95 的置信区间为

$$[\bar{X} - \mu_{\alpha/2}/\sqrt{n}, \bar{X} + \mu_{\alpha/2}/\sqrt{n}]$$

其区间长度为 $2\mu_{\alpha/2}/\sqrt{n}$, 它仅依赖于样本容量 n 且与样本取值无关. 又

$$2\mu_{\alpha/2}/\sqrt{n} \leq 1.2 \Rightarrow n \geq (2/1.2)^2 \mu_{\alpha/2}^2 = 10.67 \approx 11.$$

即样本容量为 11 时, 使得 μ 的置信度为 0.95 的置信区间长度不超过 1.2.

单个正态总体参数—— σ 未知时 μ 的置信区间

按照枢轴变量法, 有:

- σ^2 可用样本方差 $S^2 = \sum_{i=1}^n (X_i - \bar{X})^2 / (n - 1)$ 估计, 枢轴变量选为 $W = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t(n - 1)$.
- 给定置信度 $1 - \alpha$, 根据 W 的分布找出临界值 a 和 b , 使得

$$\Pr[a < W < b] = 1 - \alpha \Rightarrow b = t_{\alpha/2}(n - 1), \quad a = -t_{\alpha/2}(n - 1).$$

- 整理可得

$$\Pr \left[\bar{X} - \frac{S}{\sqrt{n}} t_{\alpha/2}(n - 1) < \mu < \bar{X} + \frac{S}{\sqrt{n}} t_{\alpha/2}(n - 1) \right] = 1 - \alpha.$$

区间估计：例 0.3

例 0.3 假设轮胎的寿命服从正态分布. 为估计某种轮胎的平均寿命, 现随机地抽取 12 只轮胎试用, 测得它们的寿命 (单位: 万千米) 如下:

4.68	4.85	4.32	4.85	4.61	5.02
5.20	4.60	4.58	4.72	4.38	4.7

试求平均寿命的 0.95 置信区间.

解答：例 0.3

题目：假设轮胎的寿命服从正态分布. 为估计某种轮胎的平均寿命, 现随机地抽取 12 只轮胎试用, 测得它们的寿命 (单位: 万千米) 如下:

4.68 4.85 4.32 4.85 4.61 5.02
5.20 4.60 4.58 4.72 4.38 4.7

试求平均寿命的 0.95 置信区间.

解答：

- 此处正态总体标准差未知, 可使用 t 分布求均值的置信区间. 本例中经计算有

$$\bar{x} = 4.709, \quad s^2 = 0.0615$$

取 $\alpha = 0.05$, 查表知 $t_{0.025}(11) = 2.2010$, 于是平均寿命的 0.95 置信区间为

$$4.7092 \pm 2.2010 \cdot \sqrt{0.0615}/\sqrt{12} = [4.5516, 4.8668]$$

单个正态总体参数 -- σ^2 的置信区间

虽然可以就 μ 是否已知分两种情况讨论 σ^2 的置信区间, 但在实际中 σ^2 未知 μ 已知的情形是极为罕见的, 所以我们只在 μ 未知的情况下讨论 σ^2 的置信区间. 按照枢轴变量法, 有:

- σ^2 可用样本方差 $S^2 = \sum_{i=1}^n (X_i - \bar{X})^2 / (n - 1)$ 估计, 枢轴变量选为 $W = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$.
- 给定置信度 $1 - \alpha$, 设临界值 a 和 b , 使得 $P[a < W < b] = 1 - \alpha$. 又根据 χ^2 分布的不对称性, 采用概率对称的区间

$$P[W \leq a] = P[W \geq b] = \frac{\alpha}{2} \Rightarrow b = \chi_{\alpha/2}^2(n-1), \quad a = \chi_{1-\alpha/2}^2(n-1).$$

- 整理可得

$$\Pr \left[\frac{(n-1)S^2}{\chi_{\alpha/2}^2(n-1)} < \sigma^2 < \frac{(n-1)S^2}{\chi_{1-\alpha/2}^2(n-1)} \right] = 1 - \alpha.$$

区间估计：例 0.4

例 0.4 某厂生产的零件质量服从正态分布 $\mathcal{N}(\mu, \sigma^2)$, 现随机地抽取 9 个零件, 测得其质量 (单位: g) 如下:

45.3 45.4 45.1 45.3 45.5 45.7 45.4 45.3 45.6

试求总体标准差 σ 的 0.95 置信区间.

解答：例 0.4

题目：某厂生产的零件质量服从正态分布 $\mathcal{N}(\mu, \sigma^2)$ ，现随机地抽取 9 个零件，测得其质量 (单位: g) 如下：

45.3 45.4 45.1 45.3 45.5 45.7 45.4 45.3 45.6

试求总体标准差 σ 的 0.95 置信区间.

解答：

- 本例中经计算有 $(n-1)S^2 = 8 \times 0.0325 = 0.26$ ，又 $\alpha = 0.05$ ，查表知 $\chi_{0.025}^2(8) = 2.1797$ ， $\chi_{0.975}^2(8) = 17.5345$ ，于是总体标准差 σ^2 的 0.95 置信区间为

$$\left[\frac{0.26}{17.5345}, \frac{0.26}{2.1797} \right] = [0.0148, 0.1193]$$

从而 σ 的 0.95 置信区间为 $[0.1218, 0.3454]$.

两个正态总体参数的置信区间

若 X_1, X_2, \dots, X_n 是来自 $\mathcal{N}(\mu_1, \sigma_1^2)$ 和 y_1, y_2, \dots, y_m 是来自 $\mathcal{N}(\mu_2, \sigma_2^2)$ 的样本, 且两个样本相互独立, 令

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n X_i, \quad \bar{y} = \frac{1}{m} \sum_{i=1}^m y_i, \quad S_1^2 = \sum_{i=1}^n \frac{(X_i - \bar{x})^2}{(n-1)}, \quad S_2^2 = \sum_{i=1}^m \frac{(y_i - \bar{y})^2}{(m-1)}.$$

下面讨论两个均值之差 $\mu_1 - \mu_2$ 和两个方差之比 σ_1^2/σ_2^2 的置信度为 $1 - \alpha$ 的区间估计.

两个正态总体参数的置信区间

- 当 σ_1^2 和 σ_2^2 已知时, 有

$$\bar{x} - \bar{y} \sim \mathcal{N} \left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m} \right) ,$$

取枢轴量为

$$W = \frac{\bar{x} - \bar{y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}} \sim \mathcal{N}(0, 1) ,$$

求解置信区间, 有

$$\Pr \left[\bar{X} - \bar{Y} - \mu_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}} < \mu_1 - \mu_2 < \bar{X} - \bar{Y} + \mu_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}} \right] = 1 - \alpha .$$

两个正态总体参数的置信区间

- 当 $\sigma_1^2 = \sigma_2^2 = \sigma^2$ 未知时, 有

$$\bar{x} - \bar{y} \sim \mathcal{N} \left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m} \right), \quad \frac{(n-1)S_1^2 + (m-1)S_2^2}{\sigma^2} \sim \chi^2(m+n-2),$$

故可以构造如下服从 t 分布的枢轴量

$$W = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_W \sqrt{\frac{1}{n} + \frac{1}{m}}} \sim t(n+m-2),$$

其中 $S_W^2 = \frac{(n-1)S_1^2 + (m-1)S_2^2}{n+m-2}$. 于是置信区间为

$$\Pr \left[-t_{\alpha/2}(n+m-2) < \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_W \sqrt{\frac{1}{n} + \frac{1}{m}}} < t_{\alpha/2}(n+m-2) \right] = 1 - \alpha.$$

两个正态总体参数的置信区间

- 求 σ_1^2/σ_2^2 的置信度为 $1 - \alpha$ 的区间估计, 由于 $\frac{(n-1)S_1^2}{\sigma_1^2} \sim \chi^2(n-1)$, $\frac{(m-1)S_2^2}{\sigma_2^2} \sim \chi^2(m-1)$, 故可构造服从 F 分布的枢轴量

$$W = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F(n-1, m-1),$$

根据 F 分布的不对称性, 采用概率对称的区间

$$\Pr[W \leq a] = \Pr[W \geq b] = \alpha/2$$

$$\Rightarrow b = F_{\frac{\alpha}{2}}(n-1, m-1), \quad a = F_{1-\alpha/2}(n-1, m-1).$$

由此可得置信区间

$$\Pr \left[\frac{S_1^2/S_2^2}{F_{\alpha/2}(n-1, m-1)} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{S_1^2/S_2^2}{F_{1-\alpha/2}(n-1, m-1)} \right] = 1 - \alpha.$$

区间估计：例 0.5

例 0.5 若从总体 $X \sim \mathcal{N}(\mu_1, \sigma_1^2)$ 和总体 $Y \sim \mathcal{N}(\mu_2, \sigma_2^2)$ 中分别抽取容量为 $n_1 = 10, n_2 = 15$ 的样本, 经计算后有 $\bar{X} = 82, S_1^2 = 56.5, \bar{Y} = 76, S_2^2 = 52.4$.

- 已知 $\sigma_1^2 = 64, \sigma_2^2 = 49$, 求 $\mu_1 - \mu_2$ 的 0.95 置信区间.
- 当 $\sigma_1^2 = \sigma_2^2 = \sigma^2$ 未知时, 求 $\mu_1 - \mu_2$ 的 0.95 置信区间.
- 求 σ_1^2/σ_2^2 的 0.95 置信区间.

解答：例 0.5

解答：

• σ_1^2 和 σ_2^2 已知时的置信区间为

$$\Pr \left[\bar{X} - \bar{Y} - \mu_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}} < \mu_1 - \mu_2 < \bar{X} - \bar{Y} + \mu_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}} \right] = 1 - \alpha.$$

查标准正态分布函数表知 $\mu_{0.975} = 1.96$, 因此 $\mu_1 - \mu_2$ 的 0.95 置信区间为

$$\left[6 - 1.96 \sqrt{\frac{64}{10} + \frac{49}{15}}, 6 + 1.96 \sqrt{\frac{64}{10} + \frac{49}{15}} \right] = [-0.0939, 12.0939]$$

• $\sigma_1^2 = \sigma_2^2$ 时, 置信区间为

$$\Pr \left[-t_{\alpha/2}(n+m-2) < \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_W \sqrt{\frac{1}{n} + \frac{1}{m}}} < t_{\alpha/2}(n+m-2) \right] = 1 - \alpha.$$

其中 $S_W^2 = \frac{(n-1)S_1^2 + (m-1)S_2^2}{n+m-2} = \frac{9 \times 56.5 + 14 \times 52.4}{23} = 54.0043$, 而 $t_{0.975}(23) = 2.0687$, 因此

$\mu_1 - \mu_2$ 的 0.95 置信区间为

$$\begin{aligned} & \left[82 - 76 - 2.0687\sqrt{54.0043}\sqrt{\frac{10+15}{10 \times 15}}, 82 - 76 + 2.0687\sqrt{54.0043}\sqrt{\frac{10+15}{10 \times 15}} \right] \\ & = [-0.2063, 12.2063]. \end{aligned}$$

• σ_1^2/σ_2^2 的置信度为 $1 - \alpha$ 的置信区间为

$$\Pr \left[\frac{S_1^2/S_2^2}{F_{\alpha/2}(n-1, m-1)} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{S_1^2/S_2^2}{F_{1-\alpha/2}(n-1, m-1)} \right] = 1 - \alpha.$$

查表得 $F_{0.975}(9, 14) = 3.21$, $F_{0.025}(9, 14) = 1/F_{0.975}(14, 9) = 1/3.8$, 因此 σ_1^2/σ_2^2 的 0.95 置信区间为

$$\left[\frac{56.5}{52.4} \cdot \frac{1}{3.21}, \frac{56.5}{52.4} \cdot 3.8 \right] = [0.3359, 4.0973]$$

非正态分布的区间估计

若总体 X 的分布未知或非正态分布, 我们可以利用集中不等式和中心极限定理给出总体期望 $\mu = \mathbb{E}[X]$ 的区间估计.

- 若 $X \in [a, b]$, 设 $\bar{X} = \sum_{i=1}^n X_i/n$, 根据集中不等式有

$$\Pr [|\mu - \bar{X}| \geq \epsilon] \leq 2 \exp(-2n\epsilon^2/(b-a)^2).$$

令

$$\alpha = 2 \exp(-2n\epsilon^2/(b-a)^2),$$

求解

$$\epsilon = \sqrt{(b-a)^2 \ln(2/\alpha)/n},$$

于是有

$$\Pr \left[\bar{X} - \sqrt{(b-a)^2 \ln(2/\alpha)/n} < \mu < \bar{X} + \sqrt{(b-a)^2 \ln(2/\alpha)/n} \right] > 1-\alpha.$$

非正态分布的区间估计

- 中心极限定理求枢轴量的近似分布, 设总体 X 的期望 $\mathbb{E}[X] = \mu$, 方差 $\mathbb{V}\mathbb{A}\mathbb{R}(X) = \sigma^2$, 有

$$W = \frac{\bar{X} - \mu}{\sigma\sqrt{n}} \xrightarrow{d} \mathcal{N}(0, 1) .$$

- 当 σ^2 已知时, 有

$$\Pr \left[-\mu_{\alpha/2} < \frac{\bar{X} - \mu}{\sigma\sqrt{n}} < \mu_{\alpha/2} \right] \approx 1 - \alpha .$$

- 当 σ^2 未知时, 用无偏样本方差 $S^2 = \sum_{i=1}^n (X_i - \bar{X})^2 / (n - 1)$ 替代 σ^2 , 于是有

$$\Pr \left[-\mu_{\alpha/2} < \frac{\bar{X} - \mu}{S/\sqrt{n}} < \mu_{\alpha/2} \right] \approx 1 - \alpha .$$

区间估计：例 0.6

例 0.6 设 X_1, X_2, \dots, X_n 是来自总体 $X \sim B(p)$ 的样本, 求 p 的置信度为 $1 - \alpha$ 的区间估计.

解答：例 0.6

解答：

- 利用集中不等式求解, 根据伯努利分布的性质有 $X_i \in \{0, 1\}$ 以及 $p = \mathbb{E}[X]$, 根据切比雪夫不等式有

$$\Pr[|\bar{X} - p| > \epsilon p] \leq 2 \exp(-np\epsilon^2/3),$$

设 $\alpha = 2 \exp(-np\epsilon^2/3)$, 于是有

$$\Pr[\bar{X} - \sqrt{3p \ln(2/\alpha)/n} < p < \bar{X} + \sqrt{3p \ln(2/\alpha)/n}] \geq 1 - \alpha,$$

- 根据伯努利分布的性质有 $p = \mathbb{E}[X]$ 以及 $\mathbb{V}\mathbb{A}\mathbb{R}(X) = p(1 - p)$, 利用中心极限定理可知

$$W = \frac{n\bar{X} - np}{\sqrt{np(1 - p)}} \xrightarrow{d} \mathcal{N}(0, 1).$$

于是有

$$\Pr\left[-\mu_{\alpha/2} < \frac{n\bar{X} - np}{\sqrt{np(1 - p)}} < \mu_{\alpha/2}\right] \approx 1 - \alpha.$$

单侧置信区间

对某些实际问题, 我们往往只关心置信区间的上限或下限

- 例如, 次品率只关心上限, 产品的寿命只关心下限

由此引入单侧置信区间及其估计.

定义 0.2 给定 $\alpha \in (0, 1)$, 设样本 X_1, \dots, X_n 的统计量 $\hat{\theta}_1 = \hat{\theta}_1(X_1, \dots, X_n)$ 满足

$$\Pr[\theta > \hat{\theta}_1] \geq 1 - \alpha,$$

则称 $(\hat{\theta}_1, +\infty)$ 为 θ 的置信度为 $1 - \alpha$ 的置信区间, 且 $\hat{\theta}_1$ 称为 **单侧置信下限**.

正态总体的单侧置信区间

对于正态总体, 可以将相关置信区间的估计都拓展到单侧置信估计.

- 若 X_1, X_2, \dots, X_n 是来自 $\mathcal{N}(\mu, \sigma^2)$ 的样本. 若 σ^2 已知, 讨论 μ 的置信度为 $1 - \alpha$ 的单侧置信下限和上限.

- 构建枢轴量为

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

- 根据单侧置信上下限的定义有

$$\Pr \left[\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < \mu_\alpha \right] = 1 - \alpha, \quad \Pr \left[\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} > -\mu_\alpha \right] = 1 - \alpha.$$

- 若 σ^2 未知, 构建枢轴量为

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

单侧置信区间估计：例 0.7

例 0.7 从一批出厂的灯泡中随机抽取 10 盏灯泡, 测试其寿命分别为: 1000, 1500, 1250, 1050, 950, 1000, 1150, 1050, 950, 1000 (单位: 小时). 假设这批灯泡的寿命服从正态分布, 求这批灯泡平均寿命的置信度为 95% 的单侧置信下限.

解答：例 0.7

题目：从一批出厂的灯泡中随机抽取 10 盏灯泡，测试其寿命分别为：1000, 1500, 1250, 1050, 950, 1000, 1150, 1050, 950, 1000 (单位：小时). 假设这批灯泡的寿命服从正态分布，求这批灯泡平均寿命的置信度为 95% 的单侧置信下限.

解答：

- 按照题意，判断此例是求 σ 未知时 μ 的单侧置信区间，枢轴变量选为 $W = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t(n-1)$. 计算样本均值及样本方差 (无偏) 分别为

$$\bar{X} = \sum_{i=1}^{10} X_i / 10 = 1090 \quad \text{和} \quad S^2 = \sum_{i=1}^{10} (X_i - \bar{X})^2 / 9 = 8800/3.$$

于是有

$$\Pr \left[\frac{\bar{X} - \mu}{S/\sqrt{10}} < t_{0.05}(9) \right] = 0.95,$$

查表可知 $t_{0.05}(9) = 1.833$ 可得

$$\mu > \bar{X} - t_{0.05}(9)S/\sqrt{10} = 1090 - \sqrt{8800/9} \times 1.833/\sqrt{10} > 1072.$$

区间估计与机器学习

- 在统计学中, 区间估计是对未知参数提供一个区间范围, 而不是一个点. 最典型的是置信区间 (Confidence Interval, CI).
- 在机器学习中, 当我们训练模型时, 不仅可以得到参数的点估计 (训练得到的权重), 也可以对预测结果或参数的不确定性进行区间估计, 从而衡量模型的不确定性.
 - 风险控制 (预测房价区间)
 - 医学诊断 (预测病人指标区间)
 - 置信学习 (Conformal Prediction)

区间估计与机器学习：线性回归的区间估计

考虑简单线性回归模型

$$y = \beta_0 + \beta_1 x + \varepsilon, \quad \varepsilon \sim \mathcal{N}(0, \sigma^2)$$

对新输入 x^* ，预测值为

$$\hat{y}^* = \hat{\beta}_0 + \hat{\beta}_1 x^*$$

则预测标准误为

$$\text{SE}(\hat{y}^*) = \sqrt{\hat{\sigma}^2 \left(1 + \frac{1}{n} + \frac{(x^* - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right)}$$

给定置信水平 $1 - \alpha$ （如 95%），则预测区间为

$$\text{CI}_{1-\alpha} = \hat{y}^* \pm t_{n-2, 1-\alpha/2} \cdot \text{SE}(\hat{y}^*)$$

线性回归的区间估计

考虑简单线性回归模型

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad \varepsilon_i \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma^2), \quad i = 1, \dots, n$$

记样本均值 $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ 以及

$$S_{xx} = \sum_{i=1}^n (x_i - \bar{x})^2.$$

参数的最小二乘估计为

$$\begin{cases} \hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{S_{xx}}, \\ \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}, \end{cases}$$

其中, $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$.

Step 1: 预测值及其方差

对新输入 x^* , 对均值响应 (mean response) 和单个新观测 (prediction) 的定义

- 均值响应的真值: $\mu(x^*) = \beta_0 + \beta_1 x^*$;
- 点预测 (拟合值): $\hat{y}^* = \hat{\beta}_0 + \hat{\beta}_1 x^*$;
- 若考虑单次真实观测: $y^* = \beta_0 + \beta_1 x^* + \varepsilon^*$, 其中 $\varepsilon^* \sim \mathcal{N}(0, \sigma^2)$ 与训练数据独立.

首先计算 \hat{y}^* 的方差 (关于样本的随机性)

$$\text{Var}(\hat{y}^*) = \text{Var}(\hat{\beta}_0 + \hat{\beta}_1 x^*) = \sigma^2 \left(\frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}} \right).$$

这个式子可由 $\text{Var}(\hat{\beta}_1) = \frac{\sigma^2}{S_{xx}}$ 、 $\text{Var}(\hat{\beta}_0) = \sigma^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}} \right)$ 和 $\text{Cov}(\hat{\beta}_0, \hat{\beta}_1) = -\frac{\bar{x}\sigma^2}{S_{xx}}$ 代入并展开得到

$$\begin{aligned}\text{Var}(\hat{\beta}_0 + \hat{\beta}_1 x^*) &= \text{Var}(\hat{\beta}_0) + x^{*2} \text{Var}(\hat{\beta}_1) + 2x^* \text{Cov}(\hat{\beta}_0, \hat{\beta}_1) \\ &= \sigma^2 \left(\frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}} \right).\end{aligned}$$

对于单次新观测 y^* ，预测误差 $y^* - \hat{y}^*$ 方差为拟合不确定性与噪声的和

$$\begin{aligned}\text{Var}(y^* - \hat{y}^*) &= \text{Var}(\varepsilon^*) + \text{Var}(\hat{y}^*) \\ &= \sigma^2 + \sigma^2 \left(\frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}} \right) \\ &= \sigma^2 \left(1 + \frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}} \right).\end{aligned}$$

因此, 若用 σ^2 表示真实误差方差, 则单次预测的标准误 (standard error) 是

$$\text{SE}_{\text{pred}}(\hat{y}^*) = \sqrt{\sigma^2 \left(1 + \frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}} \right)} .$$

Step 2: 用样本估计并构造 t 区间

在实际中 σ^2 未知, 用残差均方估计

$$\hat{\sigma}^2 = \frac{1}{n-2} \sum_{i=1}^n (y_i - \hat{y}_i)^2 .$$

常见结论 (线性回归理论) 给出

$$\frac{y^* - \hat{y}^*}{\hat{\sigma} \sqrt{1 + \frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}}}} \sim t_{n-2} ,$$

即该量服从自由度为 $n - 2$ 的 Student- t 分布（这是因为分子是均值为 0 的正态变量，分母为独立的 $\hat{\sigma}$ 的缩放，且 $(n - 2)\hat{\sigma}^2/\sigma^2 \sim \chi_{n-2}^2$ ）。

因此给定置信水平 $1 - \alpha$ ，单次观测的预测区间（prediction interval）为

$$\text{CI}_{1-\alpha}^{\text{pred}}(y^*) = \hat{y}^* \pm t_{n-2, 1-\alpha/2} \hat{\sigma} \sqrt{1 + \frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}}}$$

注意：若关注的是均值响应 $\mu(x^*)$ （不包括随机噪声 ε^* ），则其标准误为

$$\text{SE}_{\text{mean}}(\hat{y}^*) = \hat{\sigma} \sqrt{\frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}}},$$

相应的均值响应的置信区间为

$$\text{CI}_{1-\alpha}^{\text{mean}}(\mu(x^*)) = \hat{y}^* \pm t_{n-2, 1-\alpha/2} \hat{\sigma} \sqrt{\frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}}}$$

两者区别仅在于预测区间多出一个“1”，表示单次观测里有观测噪声。

区间估计与机器学习：逻辑回归的区间估计

逻辑回归模型

$$P(y = 1 \mid x; \beta) = \sigma(\beta_0 + \beta_1 x), \quad \sigma(z) = \frac{1}{1 + e^{-z}}$$

假设已经训练得到点估计 $\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1)$, 其估计协方差矩阵为 $\hat{\Sigma}$ (通常由费舍尔信息矩阵的逆得到). 则参数 β_j 的置信区间为

$$\text{CI}_{1-\alpha}(\beta_j) = \hat{\beta}_j \pm z_{1-\alpha/2} \cdot \sqrt{\hat{\Sigma}_{jj}}, \quad j = 0, 1$$

其中 $z_{1-\alpha/2}$ 为标准正态分布临界值.

对新输入 x^* 的预测概率的近似区间可以通过 delta 方法得到

$$\text{SE}(\hat{p}^*) \approx \sqrt{\nabla_{\beta} \sigma(\hat{\beta}_0 + \hat{\beta}_1 x^*)^{\top} \hat{\Sigma} \nabla_{\beta} \sigma(\hat{\beta}_0 + \hat{\beta}_1 x^*)}$$

预测概率区间为

$$\text{CI}_{1-\alpha}(\hat{p}^*) = \hat{p}^* \pm z_{1-\alpha/2} \cdot \text{SE}(\hat{p}^*)$$

神经网络中的区间估计

考虑单隐藏层神经网络回归

$$\hat{y}_i = f(x_i; \Theta), \quad \Theta = \{W_1, b_1, W_2, b_2\}$$

假设输出服从正态分布 $y_i \sim \mathcal{N}(\hat{y}_i, \sigma^2)$.

对新输入 x^* 的预测值为 $\hat{y}^* = f(x^*; \hat{\Theta})$, 其预测不确定性可以通过蒙特卡洛 dropout 或贝叶斯近似得到标准误 $\text{SE}(\hat{y}^*)$, 然后构造预测区间

$$\text{CI}_{1-\alpha}(\hat{y}^*) = \hat{y}^* \pm z_{1-\alpha/2} \cdot \text{SE}(\hat{y}^*)$$

神经网络中的区间估计

考虑单隐藏层神经网络回归

$$\hat{y}_i = f(x_i; \Theta), \quad \Theta = \{W_1, b_1, W_2, b_2\}$$

假设输出服从正态分布 $y_i \sim \mathcal{N}(\hat{y}_i, \sigma^2)$.

对新输入 x^* 的预测值为 $\hat{y}^* = f(x^*; \hat{\Theta})$, 其预测不确定性可以通过蒙特卡洛 dropout 或贝叶斯近似得到标准误 $\text{SE}(\hat{y}^*)$, 然后构造预测区间

$$\text{CI}_{1-\alpha}(\hat{y}^*) = \hat{y}^* \pm z_{1-\alpha/2} \cdot \text{SE}(\hat{y}^*)$$

例如, 通过 MC dropout 对同一个 x^* 进行 T 次前向传播, 得到预测样本 $\{\hat{y}_1, \dots, \hat{y}_T\}$, 则

$$\text{SE}(\hat{y}^*) \approx \sqrt{\frac{1}{T} \sum_{t=1}^T (\hat{y}_t - \bar{\hat{y}})^2} \quad \text{with} \quad \bar{\hat{y}} = \frac{1}{T} \sum_{t=1}^T \hat{y}_t$$

区间估计与机器学习: 小结

区间估计为机器学习模型提供了对不确定性的量化, 能够在决策中提供更可靠的信息.

- 线性回归: 可以直接利用残差和 t 分布构造预测区间.
- 逻辑回归: 可以利用参数协方差矩阵和 **delta** 方法近似得到概率预测区间.
- 神经网络: 通过蒙特卡洛 **dropout** 或贝叶斯神经网络获得预测分布, 从而计算预测区间.