概率论与数理统计

第七章

韩潇
xhan011@ustc.edu.cn

第七章：区间估计

大纲:

- 基本概念
- 枢轴变量法
- 大样本方法
- 自助法置信区间
- 置信限

7.1 基本概念

点估计的缺点：无法估计它与真实参数之间的误差有多大→能否把未知参数估计在一个区间内.

例如：把一个大一新生的年龄估计在17-19岁之间，中午食堂吃饭的费用在6-15元之间等等. 所以区间估计是一种常见的估计形式，其优点是把可能的误差用醒目的形式标出来了，即给出了参数的可能值，从而为决策提供了充分的依据

7.1 基本概念

- 设$X_1,\ldots,X_n$是从总体$F_\theta,\theta\in\Theta\subset \mathrm{R}$中抽取的一个简单随机样本，所谓$\theta$的区间估计，就是寻找两个统计量 $\hat{\theta}_1(X_1,\ldots,X_n)<\hat{\theta}_2(X_1,\ldots,X_n)$，构造区间估计$[\hat{\theta}_1,\hat{\theta}_2]$，在获得样本后，把$\theta$估计在区间$[\hat{\theta}_1,\hat{\theta}_2]$中. 由于统计量是随机变量，因此所构造的区间能否罩住参数θ?这里有个概率问题，我们希望：

➢可靠度要尽可能高，即概率

$$\mathrm{P}_\theta\left(\hat{\theta}_1(X_1,\ldots,X_n)\le\theta\le\hat{\theta}_2(X_1,\ldots,X_n)\right)$$要尽可能大

➢估计的精度要尽可能高，即区间长度尽可能小.

7.1 基本概念

定义 7.1 置信区间和置信系数

设 $(X_1, \ldots, X_n)$ 是从总体中抽取的一个简单随机样本, $\theta \in \Theta \subset \mathbb{R}$ 为未知参数. $\hat{\theta}_1(X_1, \ldots, X_n) < \hat{\theta}_2(X_1, \ldots, X_n)$ 为两个统计量, 给定一个小的正数 $\alpha \in (0,1)$, 如果

$$\mathbb{P}_\theta(\hat{\theta}_1(X_1, \ldots, X_n) \leq \theta \leq \hat{\theta}_2(X_1, \ldots, X_n)) = 1 - \alpha, \ \forall \theta \in \Theta, \tag{7.1}$$

则称区间 $[\hat{\theta}_1, \hat{\theta}_2]$ 为参数 θ 的置信区间 (Confidence interval) 估计, 置信系数 为 $1 - \alpha$. ♣

- 在实际中, 达到指定置信系数的置信区间可能有多个, 因此先寻找满足可靠性要求的区间, 然后再在满足要求的区间中寻找精度最高的区间

- 此外, 还要在最好的区间与计算方便性之间进行权衡

7.2 枢轴变量法

如何找到一个置信区间估计？直观上未知参数θ在优良的点估计附近，所以区间估计应该以这个优良的点估计为中心，向左和向右扩展一点

例：正态总体$N(\mu,\sigma^2)$，σ^2已知，如何求μ的区间估计？由于$\bar{X}$是μ的优良点估计(MVUE)，所以置信区间以$\bar{X}$中心向两边延伸. 注意到

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

由正态密度的对称性，一个合理的置信区间应该有形式$[\bar{X}-d,\bar{X}+d]$, 其中d是一个适当的常数, 称为误差界限. 换言之，在指定置信系数下，这种形式的置信区间长度最短.

7.2 枢轴变量法

如果要求置信水平为$1-\alpha$, 就要求

$$\mathbb{P}_\mu(\overline{X}-d \le \mu \le \overline{X}+d) = 1-\alpha$$
$$\Leftrightarrow \mathbb{P}_\mu\left(|\sqrt{n}(\overline{X}-\mu)/\sigma| \le \frac{\sqrt{n}d}{\sigma}\right) = 1-\alpha$$
$$\Leftrightarrow \Phi\left(\frac{\sqrt{n}d}{\sigma}\right) = 1-\frac{\alpha}{2}.$$

根据分布函数上α分位点的定义，记$u_{\alpha/2}$为方程

$$\Phi(x) = 1-\alpha/2$$

的解．则易得$\frac{\sqrt{n}d}{\sigma} = u_{\alpha/2}$.

7.2 枢轴变量法

所以 μ 的置信系数为 $1-\alpha$ 的置信区间为

$$\left[\overline{x}-\frac{\sigma}{\sqrt{n}}u_{\alpha/2},\overline{x}+\frac{\sigma}{\sqrt{n}}u_{\alpha/2}\right]\equiv\overline{x}\pm\frac{\sigma}{\sqrt{n}}u_{\alpha/2}. \tag{7.3}$$

由此，我们可以总结出寻找区间估计的一种办法-“枢轴变量法”：假设我们感兴趣的是参数θ，

1. 找一个θ的良好点估计$T(\boldsymbol{X})$，一般为θ的最大似然估计.

2. 构造一个函数$S(T,\theta,U(X))$称为枢轴变量，使得它的分布完全已知，注意枢轴变量仅是T,θ,U的函数，不能包含其他未知参数.

7.2 枢轴变量法

3. 枢轴变量必须满足如下条件：对$\forall a < b$不等式$a \le S(T,\theta,U) \le b$能改写为等价形式$A \le \theta \le B$. 其中$A, B$只能与$T(\boldsymbol{X}), U(X), a, b$有关, 与$\theta$无关
4. 取分布F的上$\alpha/2$分位点$w_{\alpha/2}$和上$1-\alpha/2$分位点 $w_{1-\alpha/2}$, 由分位点定义, 有

$$P\left(w_{1-\frac{\alpha}{2}} \le S(T,\theta,U) \le w_{\frac{\alpha}{2}}\right) = 1-\alpha,$$

根据第 3 步, 不等式 $w_{1-\frac{\alpha}{2}} \le S(T,\theta,U) \le w_{\frac{\alpha}{2}}$可以改写为$A \le \theta \le B$的形式,其中A, B 是统计量. 由置信区间的定义, 就是θ的置信系数为$1-\alpha$的置信区间.

7.2 枢轴变量法

例 7.2 设 $\mathbf{x} = (x_1, \ldots, x_n)$ 是从正态总体 $N(\mu, \sigma^2)$ 中抽取的一个样本, 参数 μ, σ^2 未知, 求 μ 的置信系数为 $1-\alpha$ 的置信区间.

例 7.3 设某小区的一号楼 1 单元 20 户居民每月煤气和水电的消费服从正态分布 $N(\mu, \sigma^2)$, 本月的平均消费为 160 元, 样本标准差为 20 元, 求居民平均付款额 μ 的置信系数为 95% 的置信区间.

上面两个例子可以看出, 在给定置信水平下, 标准差越大, 精度越低； 样本量越大, 精度越高, 与我们的常识吻合.

7.2 枢轴变量法

- 当样本量n很大时，由中心极限定理，不论总体是什么分布，只要二阶矩存在，则$\frac{\sqrt{n}(\bar{X}-\mu)}{\hat{\sigma}}$近似服从标准正态分布，这里$\hat{\sigma}$为总体标准差的相合估计.
- 关于总体均值μ的置信系数为$1-\alpha$的置信区间近似为

$$\mu \in \bar{x} \pm \frac{\hat{\sigma}}{\sqrt{n}} u_{\alpha/2}$$

总结：总体均值μ的$1-\alpha$置信区间为$\bar{x} \pm d$，其中

$$d=\begin{cases} \frac{\sigma}{\sqrt{n}} u_{\frac{\alpha}{2}}, & \text{总体为正态}, \sigma^2\text{已知} \\ \frac{s}{\sqrt{n}} t_{n-1}(\frac{\alpha}{2}), & \text{总体为正态}, \sigma^2\text{未知} \\ \frac{\hat{\sigma}}{\sqrt{n}} u_{\alpha/2}, & n>30, \sigma^2\text{未知} \end{cases}$$

7.2 枢轴变量法

例 7.4 假设国庆期间大学生的娱乐支出服从正态分布 $N(\mu,\sigma^2)$. 如果希望以 90% 的置信水平将平均娱乐支出估计在 100 元以内 (即误差边界为 100 元), 试在下述情形下计算至少要调查多少位大学生? (1) $\sigma = 200$ 元; (2) 根据经验, 大学生娱乐支出在 100 元到 1700 元之间.

例 7.5 考虑均值方差都未知时**正态总体中方差 σ^2 的置信区间**估计.

7.2 枢轴变量法

例 7.6 设有两个独立正态总体, 分别服从 $N(\mu_1,\sigma_1^2)$ 和 $N(\mu_2,\sigma_2^2)$, 其中 σ_1^2,σ_2^2 都已知. 求均值差 $\mu_2-\mu_1$ 的置信系数为 $1-\alpha$ 的置信区间. 如果 σ_1^2,σ_2^2 都未知时如何给出置信区间?

- 设X,Y的样本数为m,n, $\bar{X}\sim N(\mu_1,\frac{\sigma_1^2}{m})$, $\bar{Y}\sim N\left(\mu_2,\frac{\sigma_2^2}{n}\right)$,则枢轴变量及其分布为

$$\frac{(\bar{Y}-\bar{X})-(\mu_2-\mu_1)}{\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}}\sim N(0,1),$$

因此$\mu_2-\mu_1$的置信系数为$1-\alpha$的置信区间为$\mu_2-\mu_1\in\bar{y}-\bar{x}\pm\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}u_{\alpha/2}$

- 方差未知时，上面的方法不适用，若样本量较大(m, n>30)，则可以利用中心极限定理得到近似置信区间(提示：利用样本方差S_1^2,S_2^2)
- 若样本量不大，且$\sigma_1^2=\sigma_2^2=\sigma^2$，则有

$$(m+n-2)S_T^2=\frac{(m-1)S_1^2+(n-1)S_2^2}{\sigma^2}\sim\chi_{m+n-2}^2$$

7.2 枢轴变量法

- 则$\frac{(\bar{Y}-\bar{X})-(\mu_2-\mu_1)}{s_T\sqrt{\frac{1}{m}+\frac{1}{n}}}\sim t_{m+n-2}$
- 因此$\mu_2-\mu_1$的置信系数为$1-\alpha$的置信区间为

$$\mu_2-\mu_1\in\bar{y}-\bar{x}\pm s_T\sqrt{\frac{1}{m}+\frac{1}{n}}t_{m+n-2}(\alpha/2)$$

- 上述称为两个正态总体的均值差的区间估计.

7.2 枢轴变量法

- 考虑两个正态总体方差比的区间估计，假设总体$X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$, 样本分别为$(x_1, \dots, x_m)$，$(y_1, \dots, y_n)$, 均值未知, 做方差比值$\sigma_1^2/\sigma_2^2$的区间估计
- 由F分布的定义，我们有$\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F_{\mathrm{m}-1,n-1}$, 为枢轴变量，因此置信系数为$1-\alpha$的置信区间为

$$\frac{\sigma_1^2}{\sigma_2^2} \in [\left(\frac{s_1^2}{s_2^2}\right) F_{n-1,m-1}\left(1-\frac{\alpha}{2}\right), \left(\frac{s_1^2}{s_2^2}\right) F_{n-1,m-1}\left(\frac{\alpha}{2}\right)]$$

7.3 大样本方法

- 构造置信区间的关键是要知道枢轴变量的分布. 大样本方法就是利用极限定理，特别是中心极限定理, 来建立枢轴变量$S(T(X),\theta,U(X))$, 使得$S(T(X),\theta,U(X))$的分布与θ无关

- 如何构造比例p的置信区间？

7.3.1 比例p的区间估计

设事件A在每次试验中发生的概率为p，作n次独立试验，以Y_n记事件A发生的次数，求p的$1-\alpha$置信区间.

事实上，当n充分大时，由中心极限定理，近似有 $\frac{Y_n-np}{\sqrt{np(1-p)}} \sim N(0,1)$, 因此 $\frac{Y_n-np}{\sqrt{np(1-p)}}$可以作为构造p的置信区间估计的枢轴变量. 由中心极限定理，得

$$P\left(-u_{\alpha/2} \le \frac{Y_n-np}{\sqrt{np(1-p)}} \le u_{\alpha/2}\right) \approx 1-\alpha$$

解不等式$-u_{\alpha/2} \le \frac{Y_n-np}{\sqrt{np(1-p)}} \le u_{\alpha/2}$，令$\hat{p}=y_n/n$, 我们有（**得分区间**）

$$p \in \frac{\hat{p}+u_{\alpha/2}^2/(2n)}{1+u_{\alpha/2}^2/n} \pm u_{\alpha/2}\frac{\sqrt{\hat{p}(1-\hat{p})/n+u_{\alpha/2}^2/(4n^2)}}{1+u_{\alpha/2}^2/n}, \tag{7.7}$$

令$\hat{q}=1-\hat{p}$, 若要求区间宽度为w, 则样本量应该满足：

$$n=\frac{u_{\alpha/2}^2\left[2\hat{p}\hat{q}-w^2+\sqrt{(2\hat{p}\hat{q})^2+(1-4\hat{p}\hat{q})w^2}\right]}{w^2}$$

例 7.7 网络钓鱼 (phishing) 是一种是通过欺骗性垃圾邮件或者网页，意图引诱人们给出敏感信息 (如密码, 银行卡号等) 的一种违法攻击方式. 2016 年发表的研究论文* 描述了一项试验, 向 320 名参与者展示一些网页, 并要求他们识别哪些是合法的, 哪些是钓鱼欺诈性的. 在研究的一个阶段, 有 157 名参与者将一个钓鱼网站误认为是安全的. 记 p 为参与试验人数中错误识别钓鱼网站的比例, 求 p 的 95% 置信区间.

解 显然 p 的点估计为 $\hat{p} = 157/320 = .491$, 根据(7.7)知 p 的一个近似 95% 置信区间为

$$\frac{(.491) + (1.96)^2/2(320)}{1 + (1.96)^2/320} \pm 1.96 \frac{\sqrt{(.491)(.509)/320 + (1.96)^2/4(320)^2}}{1 + (1.96)^2/320}$$

$$= .491 \pm .054 = (.437, .545).$$

因此, 在 95% 置信水平下, 我们得出在试验设置下有 43.7% 到 54.5% 之间的参与者不能识别该钓鱼网站. □

7.3.2 一般总体均值μ的置信区间

设$(X_1, X_2, \ldots, X_n)$是从总体X中抽取的一个简单随机样本，$\mathrm{E}(X) = \mu$为感兴趣的参数，同时，总体的分布未知，如何构造μ的置信区间？

若$E(X^2) < \infty$，根据大数定律，样本标准差S是σ的一个相合估计.再由中心极限定理，近似有

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

因此 $\frac{\sqrt{\mathrm{n}}(\bar{X} - \mu)}{S}$可以作为$\mu$近似的枢轴变量，不难得到$\mu$的置信水平近似为$1 - \alpha$置信区间为

$$\mu \in \bar{x} \pm \frac{s}{\sqrt{n}} u_{\alpha/2}$$

7.3.2 一般总体均值μ的置信区间

例 7.9 设 $x_1, \ldots, x_n$ 是从 Poisson 总体 $P(\lambda)$ 中抽取的一个简单随机样本, 求 λ 的区间估计.

例 7.10 设 $x_1, \ldots, x_n$ 是从指数总体 $Exp(\lambda)$ 中抽取的一个简单随机样本, 求均值 $\theta = \lambda^{-1}$ 的 $1-\alpha$ 置信区间.

7.4 自助法置信区间

如果总体分布不是正态分布且样本量比较小，如何构建总体均值的置信区间?

自助法(Bootstrap method)-仅仅通过使用已有的样本数据而不对总体的分布做任何假设，给出置信区间

- 将手头的样本当成总体-因为样本在某种意义上反映了总体的信息

7.4 自助法置信区间

标准自助法步骤:

假设我们在当前样本观测值 $x_1, x_2, \ldots, x_n$ 下想得到统计量 $\hat{\theta}$ 的自助分布, 则

(1) 从 $x_1, x_2, \ldots, x_n$ 中有放回地抽取一组样本量同样为 n 的样本, 记为 $x_1^*, x_2^*, \ldots, x_n^*$, 称为一个自助样本.

(2) 基于自助样本 $x_1^*, x_2^*, \ldots, x_n^*$ 计算统计量 $\hat{\theta}$ 的值, 称为 $\hat{\theta}$ 的一个自助版本.

(3) 重复上述 (1)-(2) 步足够多次, 比如 B 次, 得到 B 个自助版本 $\hat{\theta}_1^*, \hat{\theta}_2^*, \ldots, \hat{\theta}_B^*$, 它们被用来近似统计量 $\hat{\theta}$ 的自助分布. 也就是说视自助版本为 $\hat{\theta}$ 的同分布复制.

7.4 自助法置信区间

构造置信区间的关键：$\hat{\theta}-\theta$可以由$\hat{\theta}^*-\hat{\theta}$来近似

定义 7.2 基本自助置信区间

设感兴趣参数 θ 的点估计为 $\hat{\theta}$, 记其 B 个自助版本统计量 $\hat{\theta}_1^*, \hat{\theta}_2^*, \ldots, \hat{\theta}_B^*$ 的 $\alpha/2$ 和 $1-\alpha/2$ 样本分位数分别为 $\hat{\theta}^*_{[(B+1)\alpha/2]}$ 和 $\hat{\theta}^*_{[(B+1)(1-\alpha/2)]}$, 则称

$$\left[2\hat{\theta}-\hat{\theta}^*_{[(B+1)(1-\alpha/2)]}, 2\hat{\theta}-\hat{\theta}^*_{[(B+1)\alpha/2]}\right]$$

为 θ 的 $1-\alpha$ 基本自助置信区间 (The Basic Bootstrap Confidence Interval). ♣

7.4 自助法置信区间

若$\hat{\theta}$是θ的(渐进)无偏估计，标准差$se(\hat{\theta})$可以得到，且枢轴变量

$$T = \frac{\hat{\theta}-\theta}{se(\hat{\theta})}$$

的(渐进)分布已知，则可以用T来构造θ的置信区间.

未知时，自助t置信区间法：

1. 使用$\hat{\theta}$的自助版本$\hat{\theta}_1^*, \hat{\theta}_2^*, \cdots, \hat{\theta}_B^*$的样本标准差估计$se(\hat{\theta})$：

$$\widehat{se}_B(\hat{\theta}) = \sqrt{\frac{1}{B-1}\sum_{b=1}^{B}(\hat{\theta}_b^* - \overline{\hat{\theta}^*})^2}, \overline{\hat{\theta}^*} = \sum_{b=1}^{B}\hat{\theta}_b^*/B$$

2. 构建一个“T类型”的变量

$$T^* = \frac{\hat{\theta}-\theta}{\widehat{se}_B(\hat{\theta})}$$

3. 再使用自助法计算出T^*的样本分位数$t_{\alpha/2}^*$和$t_{1-\alpha/2}^*$，从而得到自助t置信区间

7.4 自助法置信区间

定义 7.3 自助 t 置信区间

称

$$[\hat{\theta} - t^*_{1-\alpha/2}\hat{se}_B(\hat{\theta}), \hat{\theta} - t^*_{\alpha/2}\hat{se}_B(\hat{\theta})]$$

为 θ 的 $1-\alpha$ 自助 t 置信区间 (The Bootstrap t Confidence Interval).

记样本值为 $x=(x_1,\dots,x_n)$, 则自助 t 置信区间的算法如下

1. 由样本观测值得到 $\hat{\theta}$.
2. 对每个再抽样, $b=1,\cdots,B$:
   (a). 从 x 中有放回的抽样得到第 b 个再抽样样本 $x^{(b)}=(x_1^{(b)},\dots,x_n^{(b)})$.
   (b). 由第 b 个再抽样样本计算 $\hat{\theta}_b^*$.
   (c). 计算 $\hat{\theta}_b^*$ 的标准差估计 $\hat{se}_R(\hat{\theta}_b^*)$.(即对每个再抽样样本 $x^{(b)}$, 生成 R 个自助版本为 $\tilde{\theta}^*$, 使用自助法估计标准差):

$$\hat{se}_R(\hat{\theta}_b^*)=\sqrt{\frac{1}{R-1}\sum_{k=1}^{R}(\tilde{\theta}_k^*-\overline{\tilde{\theta}^*})^2}, \tag{7.17}$$

   其中 $\overline{\tilde{\theta}^*}=\sum_b \tilde{\theta}_b^*/R$. 注意此步嵌套使用了自助法.
   (d). 计算第 b 个重复下的"t" 统计量: $t^{(b)}=(\hat{\theta}_b^*-\hat{\theta})/\hat{se}_R(\hat{\theta}_b^*)$
3. 重复样本 $t^{(1)},\dots,t^{(B)}$ 的分布作为推断分布. 找出样本分位数 $t_{\alpha/2}^*$ 和 $t_{1-\alpha/2}^*$.
4. 计算$\hat{se}_B(\hat{\theta})$, 即自助版本 $\{\hat{\theta}_b^*\}$ 的样本标准差.
5. 计算置信界 $[\hat{\theta}-t_{1-\alpha/2}^*\hat{se}_B(\hat{\theta}),\hat{\theta}-t_{\alpha/2}^*\hat{se}_B(\hat{\theta})]$

7.5 置信限

在实际问题中，有时候我们只对参数θ一侧的界限感兴趣：

- 空气颗粒物的上限不超过某个值$(\theta \leq \bar{\theta})$
- 对于小区推出的某项惠民措施，我们仅关心支持该措施业主的百分率下限$(\theta \geq \underline{\theta})$

7.5 置信限

定义7.4

设$(X_1,X_2,\dots,X_n)$是从总体$F(x,\theta)$中抽取的一个简单随机样本, $\bar{\theta}=\bar{\theta}(X_1,X_2,\dots,X_n)$, $\underline{\theta}=\underline{\theta}(X_1,X_2,\dots,X_n)$为两个统计量.

1. 若对θ的一切可取的值，有

$$P\left(\bar{\theta}(X_1,X_2,\dots,X_n)\geq\theta\right)=1-\alpha \tag{7.18}$$

则称$\bar{\theta}$为θ的一个置信系数为$1-\alpha$的置信上限;

2. 若对θ的一切可取的值，有

$$P\left(\underline{\theta}(X_1,X_2,\dots,X_n)\leq\theta\right)=1-\alpha \tag{7.19}$$

则称$\underline{\theta}$为θ的一个置信系数为$1-\alpha$的置信下限;

7.5 置信限

- 置信上限和置信下限是一种特殊的置信区间，其一端为∞或−∞，因此前面求区间估计的方法，可以平移到此处.
- 求正态分布总体$N(\mu,\sigma^2)$的均值μ的置信上限，当σ^2已知时，$\sqrt{n}(\bar{X}-\mu)/\sigma$是枢轴变量，分布为标准正态，故

$$P\left(\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma} \geq -u_\alpha\right) = 1-\alpha \Rightarrow P\left(\bar{X}+\frac{\sigma}{\sqrt{n}}u_\alpha \geq \mu\right) = 1-\alpha$$

参考(7. 18)，我们有$\bar{x}+\frac{\sigma}{\sqrt{n}}u_\alpha$是$\mu$的$1-\alpha$的置信上限

- 同理，μ的$1-\alpha$的置信下限为$\bar{x}-\frac{\sigma}{\sqrt{n}}u_\alpha$.
- σ^2未知时, μ的$1-\alpha$的置信上限和置信下限分别为$\bar{x}+\frac{s}{\sqrt{n}}t_{n-1}(\alpha)$和$\bar{x}-\frac{s}{\sqrt{n}}t_{n-1}(\alpha)$
- 正态总体中σ^2的$1-\alpha$的置信上限和置信下限分别为$(n-1)s^2/\chi^2_{n-1}(1-\alpha)$和$(n-1)s^2/\chi^2_{n-1}(\alpha)$

7.5 置信限

例 7.13 设某地区 12 个测量点测得某时刻空气质量指数 (Air Quality Index, AQI)($\mu g/m^3$) 为

$$29,\ 37,\ 47,\ 57,\ 53,\ 60,\ 54,\ 51,\ 46,\ 44,\ 33,\ 32.$$

设空气质量指数服从正态分布, 求该时刻某地区空气质量指数的 90% 的置信上限. 能否说该时刻某地区的空气为优? (空气质量指数在 $50\mu g/m^3$ 以下为优)

Thank you!