

1 主题内容与适用范围

本标准规定了常用的数理统计术语。

本标准适用于各类标准与技术文件中涉及的数理统计术语,对各类研究技术报告和著作中涉及的数理统计术语也应参照使用。

2 概率论术语

2.1 概率 probability

度量一随机事件发生可能性大小的实数,其值介于 0 与 1 之间。

注:一随机事件的概率可看作在相同条件下重复试验时,该事件发生的频率的稳定值,也可看作对事件发生的相信程度。

2.2 〔一维〕随机变量〔univariate〕random variable, variate

取值随试验结果而定,且有一定的概率分布的变量。

2.3 〔一维〕概率分布〔univariate〕probability distribution

给出一个随机变量取任何给定值或取值于任何给定集合的概率的函数,随机变量在其整个变化区域取值的概率为 1。

2.4 〔一维〕分布函数〔univariate〕distribution function

随机变量 X 小于或等于实数 x 的概率,它是 x 的函数:

$$F(x) = P(X \leq x)$$

注: $F(x)$ 是右连续的。有时也把 $F(x) = P(X < x)$ 定义为分布函数,此时 $F(x)$ 是左连续的。

2.5 连续随机变量及〔概率〕密度函数 continuous random variable and〔probability〕density function

如果随机变量的分布函数 $F(x)$ 可表示为一非负函数 $f(x)$ 的积分:

$$F(x) = \int_{-\infty}^x f(t) dt,$$

则称该随机变量为连续随机变量, $f(x)$ 称为它的〔概率〕密度函数。

任何密度函数 $f(x)$ 必满足

$$\int_{-\infty}^{\infty} f(x) dx = 1.$$

注:连续随机变量的概率分布称为连续分布。

2.6 离散随机变量及概率函数 discrete random variable and probability function

只能取有限或可列个值 (x_1, x_2, \dots) 的随机变量 X 称为离散随机变量。给出 X 取每一可能值 x_i 的

概率的函数称为概率函数:

$$p_i = P(X = x_i), i = 1, 2, \dots$$

任何概率函数必满足 $\sum_i p_i = 1$ 。

注: 离散随机变量的概率分布称为离散分布。

2.7 k 维随机变量 k -dimensional random variable

取值随试验结果而定, 且有一定的 k 维概率分布的 k 维向量:

$$X = (X_1, X_2, \dots, X_k)。$$

2.8 k 维概率分布 k -dimensional probability distribution

给出一个 k 维随机变量取值于 k 维空间中任何给定集合的概率的函数。 k 维 ($k \geq 2$) 概率分布也称为 k 个随机变量的联合分布。

注: ① k 维空间上的实值函数

$$F(x_1, x_2, \dots, x_k) = P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_k \leq x_k)$$

称为 k 维随机变量 $X = (X_1, X_2, \dots, X_k)$ 的联合分布函数。

② 如果 k 维随机变量的分布函数 $F(x_1, x_2, \dots, x_k)$ 可表示为一非负函数 $f(x_1, x_2, \dots, x_k)$ 的积分:

$$F(x_1, x_2, \dots, x_k) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_k} f(t_1, t_2, \dots, t_k) dt_1 \dots dt_k,$$

则称该 k 维随机变量为 k 维连续随机变量, $f(x_1, x_2, \dots, x_k)$ 称为它的〔概率〕密度函数。 k 维连续随机变量的分布称为 k 维连续分布。

只能取有限或可列组值 $(x_{11}, x_{12}, \dots, x_{1k})$ 的 k 维随机变量称为 k 维离散随机变量, 给出它取每一组可能值 $(x_{11}, x_{12}, \dots, x_{1k})$ 的概率的函数称为 k 维概率函数。

$$p_i = P(X_1 = x_{11}, \dots, X_k = x_{1k}), i = 1, 2, \dots$$

k 维离散随机变量的分布称为 k 维离散分布。

2.9 边缘分布 marginal distribution

k 维随机变量的 p 个分量的联合分布, 称为该 k 维随机变量的分布的 p 维边缘分布。

例: 三维随机变量 (X, Y, Z) 含有:

——三个二维边缘分布, 即 (X, Y) , (X, Z) 和 (Y, Z) 的分布;

——三个一维边缘分布, 即 X, Y 和 Z 的分布。

2.10 条件分布 conditional distribution

k 维随机变量的 p 个分量在另外 $k-p$ 个分量取给定值条件下的联合概率分布。

2.11 独立 independence

若两组随机变量有联合概率分布, 其任何一组的条件分布都不随另一组的取值而变化, 则称它们是独立的。否则称它们是相依的 (dependence)。

注: ① 若两组随机变量 (X_1, \dots, X_m) 和 (Y_1, \dots, Y_n) 是独立的, 则对任意 (x_1, \dots, x_m) 和 (y_1, \dots, y_n) 都有

$$F(x_1, \dots, x_m; y_1, \dots, y_n) = F_1(x_1, \dots, x_m)F_2(y_1, \dots, y_n),$$

式中 $F_1(x_1, \dots, x_m)$ 与 $F_2(y_1, \dots, y_n)$ 分别为 (X_1, \dots, X_m) 和 (Y_1, \dots, Y_n) 的分布函数, $F(x_1, \dots, x_m; y_1, \dots, y_n)$ 为它们的联合分布函数。

对密度函数和概率函数, 类似公式也成立。

② 上述概念可以推广到 n 组随机变量的情形。

2.12 分位数 quantile

对随机变量 X , 满足条件 $P(X < x_p) \leq p$ 和 $P(X \leq x_p) \geq p$ 的实数 x_p 称为 X 的或其分布的 p 分位数。

注: ① p 分位数可以不唯一。

② $x_{0.75}$ 与 $x_{0.25}$ 分别称为上、下四分位数 (quartile)。

2.13 中位数 median

随机变量或它的概率分布的 0.5 分位数。

2.14 众数 mode

密度函数或概率函数达到极大值的点。

注：若众数只有一个，则该概率分布称为单峰的。

2.15 期望 expectation

a. 对以概率 p_i 取值 x_i 的离散随机变量 X ，其期望定义为：

$$E(X) = \sum_i p_i x_i,$$

式中 Σ 是对 X 的所有可能值 x_i 求和。

b. 对密度函数为 $f(x)$ 的连续随机变量 X ，其期望定义为：

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx.$$

随机变量的期望也称为它的概率分布的期望。

同义词：均值 mean

注：上述求和与积分都要求绝对收敛。

2.16 条件期望 conditional expectation

随机变量的条件分布的期望。

2.17 中心化随机变量 central random variable

随机变量 X 与其期望之差： $X - E(X)$ 。

注：中心化的目的是为使随机变量的期望为零。

2.18 方差 variance

随机变量 X 的方差定义为：

$$V(X) = E[(X - E(X))^2].$$

它亦称为 X 的分布的方差。

2.19 标准差 standard deviation

方差的正平方根 $\sqrt{V(X)}$ 。

2.20 变异系数 coefficient of variation

标准差与期望的绝对值之比 $\sqrt{V(X)} / |E(X)|$ 。

2.21 标准化随机变量 standardized random variable

中心化随机变量 $X - E(X)$ 与其标准差之商：

$$\frac{X - E(X)}{\sqrt{V(X)}}.$$

注：标准化的目的是为使随机变量的期望为 0，方差为 1。

2.22 矩 moment

随机变量幂函数的期望。

2.23 原点矩 moment about the origin

随机变量 X 的 q 阶原点矩 (q 是正整数) 是指 X 的 q 次幂的期望：

$$E(X^q).$$

注：一阶原点矩即是期望。

2.24 中心矩 central moment

随机变量 X 的 q 阶中心矩 (q 是正整数) 是指：

$$E[(X - E(X))^q].$$

注：二阶中心矩即是方差。

2.25 联合矩 joint moment

二维随机变量 (X, Y) 的 (q, s) 阶联合原点矩(q, s 皆为正整数)是指:

$$E(X^q Y^s).$$

二维随机变量 (X, Y) 的 (q, s) 阶联合中心矩是指:

$$E[(X - E(X))^q (Y - E(Y))^s].$$

注:可类似定义多维随机变量的联合矩。

2.26 协方差 covariance

二维随机变量 (X, Y) 的协方差是指其 $(1, 1)$ 阶联合中心矩:

$$C_{ov}(X, Y) = E[(X - E(X))(Y - E(Y))].$$

2.27 相关系数 correlation coefficient

二维随机变量 (X, Y) 的相关系数,是指其协方差与 X, Y 的标准差乘积之比:

$$\rho = \frac{C_{ov}(X, Y)}{\sqrt{V(X)V(Y)}}.$$

当 $\rho=0$ 时,称此两随机变量不相关。

2.28 绝对矩 absolute moment

在2.23条至2.25条关于矩的定义中 $X, X - E(X), Y, Y - E(Y)$ 等量分别以其绝对值 $|X|, |X - E(X)|, |Y|, |Y - E(Y)|$ 等代替,这样定义的矩称为相应的绝对矩。

注:对于绝对矩, q 和 s 为正数,但不必为整数。

例1: X 的 q 阶绝对中心矩是指:

$$E[|X - E(X)|^q].$$

例2: (X, Y) 的 (q, s) 阶联合绝对原点矩是指:

$$E(|X|^q |Y|^s).$$

2.29 峰度 kurtosis

随机变量的概率分布的峰度是相应标准化随机变量的4阶原点矩减3,它描述了该标准化随机变量的分布的陡峭程度。

例:正态分布的峰度为0。

2.30 偏度 skewness

随机变量的概率分布的偏度是相应标准化随机变量的3阶原点矩,它描述了该分布的非对称程度。

注:偏度为零时,分布称为对称的;偏度不为0时,分布称为偏斜的。当偏度为正时,称为右偏;偏度为负时,称为左偏。

例:正态分布的偏度为0。

2.31 回归函数,回归方程与回归系数 regression function, regression equation and regression coefficient

若 Y 是一随机变量, X 是一 k 维随机变量,则给定 $X = (x_1, x_2, \dots, x_k)$ 时 Y 的条件期望 $f(x_1, x_2, \dots, x_k)$ 作为 x_1, x_2, \dots, x_k 的函数,称为 Y 对 X 的回归函数。此时 $y = f(x_1, x_2, \dots, x_k)$ 称为回归方程,回归函数的图形称为回归曲面。当回归曲面是一超平面

$$y = b_0 + b_1 x_1 + b_2 x_2 + \dots + b_k x_k$$

时,回归称为线性的,此时 b_i 称为 Y 对 X_i 的偏回归系数($i=1, 2, \dots, k$)。

当 X 是一维随机变量时, Y 对 X 的回归函数 $f(x)$ 的图形称为回归曲线,当 Y 对 X 的回归曲线是一直线 $y = a + bx$ 时,称为回归直线,其斜率 b 称为 Y 对 X 的回归系数。

2.32 均匀分布 uniform distribution

一种连续概率分布,其密度函数为

$$f(x) = \begin{cases} \frac{1}{b-a}, & \text{当 } a \leq x \leq b; \\ 0, & \text{其他。} \end{cases}$$

2.33 正态分布 normal distribution

一种连续概率分布,其密度函数为

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right], -\infty < x < \infty。$$

式中 $-\infty < \mu < \infty, 0 < \sigma < \infty$ 。

注: μ 和 σ 分别为正态分布的期望与标准差。

2.34 标准正态分布 standardized normal distribution

标准化正态随机变量的分布,其密度函数为:

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right), -\infty < x < \infty。$$

2.35 χ^2 分布 chi-square distribution

一种连续概率分布,其密度函数为:

$$f(x) = \frac{x^{\nu/2-1}}{2^{\nu/2}\Gamma(\nu/2)} \exp\left(-\frac{x}{2}\right), x \geq 0,$$

式中 ν 为正整数,称为分布的自由度。

注: 自由度为 ν 时的 χ^2 分布是 ν 个相互独立的标准正态随机变量平方和的分布。

2.36 t 分布 t -distribution, Student's distribution

一种连续概率分布,其密度函数为:

$$f(x) = \frac{\Gamma[(\nu+1)/2]}{\sqrt{\pi\nu}\Gamma(\nu/2)} \left(1 + \frac{x^2}{\nu}\right)^{-\frac{(\nu+1)}{2}}, -\infty < x < \infty,$$

式中 $\nu > 0$ 称为分布的自由度。

注: 自由度 ν 为正整数时的 t 分布是两个独立随机变量之商的分布。分子是标准正态随机变量,分母是自由度为 ν 的 χ^2 分布随机变量被其自由度除所得商的正平方根。

2.37 F 分布 F -distribution

一种连续概率分布,其密度函数为:

$$f(x) = \frac{\Gamma[(\nu_1 + \nu_2)]}{\Gamma(\nu_1/2)\Gamma(\nu_2/2)} \nu_1^{\nu_1/2} \nu_2^{\nu_2/2} \frac{x^{(\nu_1/2)-1}}{(\nu_1 x + \nu_2)^{(\nu_1 + \nu_2)/2}}, x > 0,$$

式中 ν_1, ν_2 为正整数,分别称为分布的第一自由度与第二自由度。

注: 自由度为 ν_1, ν_2 的 F 分布是两个独立的随机变量之商的分布,分子和分母都是 χ^2 分布随机变量除以各自的自由度 ν_1, ν_2 。

2.38 对数正态分布 log-normal distribution

一种连续概率分布,其密度函数为:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma x} \exp\left[-\frac{1}{2}\left(\frac{\ln x - \mu}{\sigma}\right)^2\right], x > 0,$$

式中 $-\infty < \mu < \infty, 0 < \sigma < \infty$ 。

注：① 若 X 是对数正态分布，则 $\ln X$ 的概率分布是期望为 μ ，标准差为 σ 的正态分布。

② $f(x)$ 中的自然对数 \ln 可用常用对数 \lg 代替，此时，

$$f(x) = \frac{0.4343}{\sqrt{2\pi\sigma x}} \exp\left[-\frac{1}{2}\left(\frac{\lg x - \mu}{\sigma}\right)^2\right], x > 0.$$

③ 上述密度函数中可用 $x-\gamma(x \geq \gamma)$ 代替 x ，从而得到推广。此时， γ 为位置参数。

2.39 指数分布 exponential distribution

一种连续概率分布，其密度函数为：

$$f(x) = \lambda e^{-\lambda x}, x > 0,$$

式中 $\lambda > 0$ 。

注：上述密度函数中可用 $x-\gamma(x \geq \gamma)$ 代替 x ，从而得到推广。此时， γ 为位置参数。

2.40 Γ 分布 gamma distribution

一种连续概率分布，其密度函数为：

$$f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} \exp\left(-\frac{x}{\beta}\right), x > 0,$$

式中 $\alpha > 0, \beta > 0$ 分别称为分布的形状参数与尺度参数，而

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx.$$

注：① 上述 Γ 分布密度函数中可用 $x-\gamma(x \geq \gamma)$ 代替 x ，从而得到推广，此时 γ 称为位置参数。

② 参数为 $\alpha = \nu/2, \beta = 2$ 的 Γ 分布，即是自由度为 ν 的 χ^2 分布。

2.41 B 分布 beta distribution

一种连续分布，其密度函数为：

$$f(x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1}, 0 < x < 1$$

式中 $\alpha > 0, \beta > 0$ 为分布的形状参数。

2.42 I 型极值分布 type I extreme value distribution, Gumbel distribution

一种连续概率分布，其分布函数为：

$$F(x) = \exp[-e^{-(x-\gamma)/\beta}], -\infty < x < \infty,$$

式中 $\beta > 0, -\infty < \gamma < \infty$ 。

2.43 II 型极值分布 type II extreme value distribution, Fréchet distribution

一种连续概率分布，其分布函数为

$$F(x) = \exp\{-[(x-\gamma)/\beta]^{-\alpha}\}, x \geq \gamma,$$

式中 $\alpha > 0, \beta > 0, -\infty < \gamma < \infty$ 。

2.44 威布尔分布 Weibull distribution, type III extreme value distribution

一种连续概率分布，其分布函数为

$$F(x) = 1 - \exp\left[-\left(\frac{x-\gamma}{\beta}\right)^\alpha\right], x \geq \gamma,$$

式中 $\alpha > 0, \beta > 0, -\infty < \gamma < \infty$ 分别称为分布的形状参数，尺度参数与位置参数。

2.45 二项分布 binomial distribution

一种离散概率分布,其概率函数为

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}, x = 0, 1, \dots, n,$$

式中 $0 < p < 1$, $\binom{n}{x} = \frac{n!}{x!(n-x)!}$ 。

2.46 负二项分布 negative binomial distribution

一种离散概率分布,其概率函数为

$$P(X = x) = \frac{c(c+1)\cdots(c+x-1)}{x!} p^c (1-p)^x, x = 0, 1, \dots,$$

式中 $c > 0$, $0 < p < 1$ 。

2.47 泊松分布 Poisson distribution

一种离散概率分布,其概率函数为

$$P(X = x) = e^{-\lambda} \frac{\lambda^x}{x!}, x = 0, 1, \dots,$$

式中 $\lambda > 0$ 。

注:泊松分布的期望和方差都是 λ 。

2.48 超几何分布 hypergeometric distribution

一种离散概率分布,其概率函数为

$$P(X = x) = \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}},$$

式中 $N, M (\leq N), n (\leq N)$ 是正整数, x 为整数,其取值范围为 $\max(0, n+M-N) \leq x \leq \min(M, n)$ 。

2.49 二维正态分布 bivariate normal distribution

一种连续二维概率分布,其密度函数为

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right) + \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right]\right\},$$

$$-\infty < x, y < \infty,$$

式中 $-\infty < \mu_x < \infty, -\infty < \mu_y < \infty, \sigma_x > 0, \sigma_y > 0, -1 < \rho < 1$ 。

注:① X 和 Y 的边缘分布都是正态分布,其期望分别为 μ_x 和 μ_y ,标准差分别为 σ_x 和 σ_y , ρ 是 X 和 Y 的相关系数。

② 这一概念可推广到 $k (k > 2)$ 维情形。

2.50 多项分布 multinomial distribution

一种离散多维概率分布,其概率函数为

$$P(X_1 = x_1, \dots, X_k = x_k) = \frac{n!}{x_1! \cdots x_k!} p_1^{x_1} \cdots p_k^{x_k},$$

$$x_i = 0, 1, \dots, n, (i = 1, \dots, k), \sum_{i=1}^k x_i = n,$$

式中 $p_i \geq 0 (i=1, 2, \dots, k)$, $\sum_{i=1}^k p_i = 1$ 。

3 基本统计术语

3.1 个体 item, individual

可以单独观测和研究的一个物体、一定量的材料或一次服务。也指表示上述物体、材料或服务的一个定量或定性的特性值。

3.2 总体 population

一个统计问题中所涉及个体的全体。

3.3 总体分布 population distribution

当个体理解为定量特性值时,总体的每一个体可看成是某一确定的随机变量的一个观测值,称这个随机变量的分布为总体分布。

3.4 特性 characteristic

所考察的定性或定量的性质或指标。

注:特性在任一特定个体上的值称为特性值。

3.5 样本 sample

按一定程序从总体中抽取的一组(一个或多个)个体(或抽样单元,见 5.2)。

注:① 样本中的每个个体有时也称为样品。

② 若样本是按某种随机方式抽取的,则样本可以看成是一组随机变量,其中每一个随机变量也称为样本分量。

3.6 抽样 sampling

从总体中抽取样本。

3.7 样本量 sample size

样本中所包含的个体(或抽样单元)的数目。

3.8 独立同分布样本 independently identically distributed sample

分量的分布与总体分布相同且各分量相互独立的样本。

注:在数理统计中,独立同分布样本通常称为简单随机样本(simple random sample)。在使用此术语时,要注意与 5.7 中的术语相区别。

3.9 观测值 observed value

作为一次观测结果而确定的特性值。

3.10 组 class

对于定量特性,将该特性的整个变化区间分成相连接而不重叠的若干小区间,这种小区间称为组。

3.11 组限 class limits

组的上、下界限。

注:应明确规定上、下限中的哪一个属于该组。

3.12 组中值 mid-point of class

组的上、下限的算术平均值。

3.13 组距 class width

组的上、下限之差。

3.14 频数 absolute frequency

多次观测中一给定事件发生的次数,或落入一特定组的观测值个数。

3.15 累积频数 cumulative absolute frequency

在定量特性情形,小于或等于某给定值,或某给定组的上限的观测值个数。

3.16 频率 relative frequency

频数与试验或观测总次数之比。

3.17 累积频率 cumulative relative frequency

累积频数与试验或观测总次数之比。

3.18 直方图 histogram

连续随机变量观测值分布状况的一种图形表示。在横坐标轴上将该随机变量的取值区间分为组，分别以各组为底作矩形，其面积等于相应组的频率(频数)。

注：以频率(频数)表示的直方图称为频率(频数)直方图。

3.19 条形图 bar chart

离散随机变量观测值分布状况的一种图形表示，在一坐标轴上点出观测值的数值，分别从这些点出发向同一方向作与该坐标轴垂直的线条，其长度等于相应的观测值的频率(频数)。

3.20 散点图 scatter diagram

两个随机变量的每一对观测值用直角坐标平面上的一个点表示所成的图形。

3.21 列联表 contingency table

观测数据按两个或更多定性特性分类时所列出的频数表。

注：对于定量特性，若将它们按其值分成组，也可列出列联表。

3.22 统计量 statistic

样本的函数，它不依赖于未知参数。

3.23 样本均值 sample mean

样本 X_1, \dots, X_n 的算术平均数：

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i.$$

3.24 次序统计量 order statistics

将样本的各分量从小到大排列成 $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ ，称 $(X_{(1)}, X_{(2)}, \dots, X_{(n)})$ 为次序统计量， $X_{(i)}$ 称为第 i 个次序统计量。

3.25 样本中位数 sample median

当样本量 n 为奇数时，样本中位数是第 $(n+1)/2$ 个次序统计量；当 n 为偶数时，是第 $n/2$ 个与第 $n/2+1$ 个次序统计量的算术平均数。

3.26 中程数 midrange

样本中最大值与最小值的算术平均数：

$$(X_{(1)} + X_{(n)})/2.$$

3.27 极差 range

样本中最大值与最小值之差：

$$X_{(n)} - X_{(1)}.$$

3.28 平均绝对差 mean deviation

样本分量与样本均值之绝对差的算术平均数：

$$\frac{1}{n} \sum_{i=1}^n |X_i - \bar{X}|.$$

3.29 样本方差 sample variance

样本分量与样本均值之差的平方和,除以样本量减1:

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2,$$

其中 $n-1$ 为自由度。

注:当涉及多个随机变量时,可用下标表明相应的随机变量,例如记 X 的方差为 S_x^2 。

3.30 样本标准差 sample standard deviation

样本方差的正平方根。

3.31 样本变异系数 sample coefficient of variation

样本标准差与样本均值的绝对值之比。

3.32 样本协方差 sample covariance

二维样本 $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$ 的样本协方差是:

$$S_{xy} = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y}).$$

3.33 样本相关系数 sample correlation coefficient

二维样本 $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$ 的样本相关系数是:

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2}}$$

3.34 经验分布 empirical distribution

对样本 X_1, X_2, \dots, X_n 的每个分量 X_i 赋予相等概率 $1/n$ 所得的概率分布。

经验分布的分布函数称为经验分布函数:

$$F_n(x) = \begin{cases} 0, & x < X_{(1)}; \\ \frac{i}{n}, & X_{(i)} \leq x < X_{(i+1)}, i = 1, 2, \dots, n-1; \\ 1, & x \geq X_{(n)} \end{cases}$$

注:①对取定的一组样本观测值 x_1, x_2, \dots, x_n , 经验分布是一个确定的离散分布。

②对任意给定的数值 x , $F_n(x)$ 是样本的函数,它是一个统计量。

3.35 样本矩 sample moment

经验分布的矩。

例1:对样本 X_1, X_2, \dots, X_n 及正整数 q , 样本 q 阶原点矩是指

$$\frac{1}{n} \sum_{i=1}^n X_i^q;$$

当 $q=1$ 时,即样本均值 \bar{X} 。

例2:对样本 X_1, X_2, \dots, X_n 及正整数 q , 样本 q 阶中心矩是指

$$\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^q,$$

当 $q=2$ 时,即样本二阶中心矩

$$\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{n-1}{n} S^2,$$

其中 S^2 是样本方差。

- 3.36 经验回归方程 empirical regression equation
根据样本,对回归方程所作的估计(见 3.40)。其图形称为经验回归曲面或经验回归曲线。
- 3.37 经验回归系数 empirical regression coefficient
经验回归方程中相应变量的系数。它是根据样本对回归系数作出的估计。
- 3.38 游程 run
在属性观测的系列中,同一属性的不间断的完整子系列。
例:在表示为“+”,“-”两种属性的以下观测系列中++-+-+---++++共有 4 个“+”游程,3 个“-”游程。
- 3.39 估计 estimation
根据样本推断总体分布的未知成分,例如参数。
- 3.40 估计量 estimator
用以估计总体分布未知量的统计量。
- 3.41 估计值 estimate
根据样本观测值,对估计量的计算结果。
- 3.42 估计量的偏倚 bias of estimator
估计量的期望与被估未知量真值之差。
- 3.43 均方误差 mean square error
估计量与被估未知量真值之差平方的期望。
注:估计量的均方误差等于估计量的方差与其偏倚的平方之和。
- 3.44 抽样误差 sampling error
由于样本的随机性而产生的误差。
- 3.45 无偏估计量 unbiased estimator
期望等于被估未知量真值的估计量。
- 3.46 标准误差 standard error
估计量的标准差。
注:标准误差通常用于估计量是无偏的或近似无偏的情形。
- 3.47 双侧置信区间 two-sided confidence interval
若 θ 是要估计的总体分布未知量, $T_1 \leq T_2$ 是两个统计量,使区间 $[T_1, T_2]$ 以一定概率包含 θ ,则称此区间是 θ 的一个双侧置信区间。 T_2 和 T_1 分别称为置信区间的上、下限。
- 3.48 单侧置信区间 one-sided confidence interval
在置信区间 $[T_1, T_2]$ 中,当上限 T_2 为 ∞ 或未知量的上限,或者当下限 T_1 为 $-\infty$ 或未知量的下限时,称该置信区间为单侧置信区间。此时,对于前者, T_1 称为置信下限;对于后者, T_2 称为置信上限。
- 3.49 置信水平 confidence level
 $[T_1, T_2]$ 是 θ 的一个双侧或单侧置信区间, $1-\alpha$ 是 0 和 1 之间的常数,若对一切 θ , 有

$$P(T_1 \leq \theta \leq T_2) \geq 1 - \alpha,$$

则称 $1-\alpha$ 为该置信区间的置信水平。

注:① 当 $P(T_1 \leq \theta \leq T_2) = 1-\alpha$ 时, $1-\alpha$ 也常称为置信系数或置信度。

② 置信水平 $1-\alpha$ 通常取接近于 1 的值,如 0.90, 0.95, 0.99 等。

3.50 统计覆盖区间 statistical coverage interval

由满足 $T_1 \leq T_2$ 的两个统计量构成的区间 $[T_1, T_2]$, 它以不低于 γ 的概率至少包含总体的一确定比例 β , 即

$$P[F(T_2) - F(T_1) \geq \beta] \geq \gamma,$$

则称此区间为总体分布 $F(x)$ 的 (β, γ) 统计覆盖区间, T_2, T_1 分别称为该统计覆盖区间的上、下限。

3.51 拟合优度 goodness of fit

观测值与事先假定的分布(模型)之间的符合程度的数值刻画。

3.52 离群值 outlier

样本中的一个或几个观测值, 它们离开其他观测值较远, 暗示它们可能来自不同的总体。

3.53 统计假设 statistical hypothesis

关于一个或多个总体分布的命题, 它可以通过样本去进行检验。

3.54 统计检验 statistical test

根据样本, 决定某个统计假设应该被拒绝或不被拒绝(接受)的方法和步骤。

3.55 原假设与备择假设 null hypothesis and alternative hypothesis

原假设 H_0 是一个特定的统计假设, 对它要作出拒绝或接受的决定。异于原假设, 且在原假设被拒绝时可能采用的统计假设称为备择假设。

例 1: 关于假设期望 μ 不小于给定值 μ_0 的检验问题可表述为:

$$H_0: \mu \geq \mu_0 \leftrightarrow H_1: \mu < \mu_0.$$

例 2: 关于假设两批产品不合格品率相等(但未知)的检验问题可表述为:

$$H_0: p_1 = p_2 \leftrightarrow H_1: p_1 \neq p_2.$$

例 3: 关于假设总体分布为正态分布(参数不确定)的检验问题, 备择假设为总体分布不是正态的。

例 4: 关于泊松分布中参数 λ 等于给定值 $\lambda_0 (> 0)$ 的检验问题可表述为:

$$H_0: \lambda = \lambda_0 \leftrightarrow H_1: \lambda \neq \lambda_0.$$

3.56 简单假设 simple hypothesis

完全确定了总体分布的统计假设。

注: 3.55 条的例 4 中的 H_0 是简单假设。

3.57 复合假设 composite hypothesis

不完全确定总体分布的统计假设。

例 1: 在正态分布 $N(\mu, \sigma^2)$ 的假定下, 当标准差 σ 已知时, 假设 $\mu = \mu_0$ 是简单假设, 而当 σ 未知时, 则是复合假设。

例 2: 在 3.55 条的例中给出的前三个统计假设都是复合假设。

3.58 检验统计量 test statistic

取值决定一个统计假设被拒绝与否的统计量。

3.59 非参数检验 non-parametric test

当一个统计假设不能用有限个参数来描述时所采用的检验方法。

例: 检验一个样本是否来自某一分布的柯尔莫哥洛夫检验。

3.60 拒绝域 rejection region

检验统计量取值的一个集合, 如果该统计量的观测值属于这个集合, 则原假设被拒绝; 否则, 原假设不被拒绝, 拒绝域也称为否定域。

3.61 临界值 critical value

拒绝域的边界值。

例：当正态分布的标准差 σ 已知时，对关于均值的检验问题：

$$H_0: \mu \geq \mu_0 \leftrightarrow H_1: \mu < \mu_0,$$

一种常用的检验方法是以样本均值 \bar{X} 作为检验统计量，拒绝域是所有小于

$$A = \mu_0 + u_\alpha \sigma / \sqrt{n}$$

的数构成的集合， A 就是临界值，此处 n 为样本量， u_α 为标准正态分布的 α 分位数。

3.62 单侧检验 one-sided test

在检验统计量是一维的情形，以小于（大于）一给定数的所有数值的集合作为拒绝域的检验。

3.63 双侧检验 two-sided test

在检验统计量是一维的情形时，以直线上一个有限区间的外部作为拒绝域的检验。

注：选择单侧检验还是双侧检验取决于备择假设。在 3.55 条的例 1 中，检验是单侧的，而例 2 的检验是双侧的。

3.64 第一类错误 type I error

原假设为真而被拒绝。又称弃真。

3.65 第一类错误概率 type I error probability

一个检验犯第一类错误的概率。

3.66 第二类错误 type II error

原假设不真但被接受。又称存伪。

3.67 第二类错误概率 type II error probability

一个检验犯第二类错误的概率。

3.68 显著性水平 significant level

检验的第一类错误概率不可超过的界限 α ， α 一般取较小的数值，如 0.10, 0.05, 0.01 等。

3.69 检验的功效 power of a test

当原假设不真时，拒绝原假设的概率。

3.70 功效函数 power function

总体参数的函数，它是当该参数为真时拒绝原假设的概率。

3.71 功效曲线 power curve

功效函数的图形表示。

例 1：图 1 是 3.61 条的例中所述检验的功效曲线。

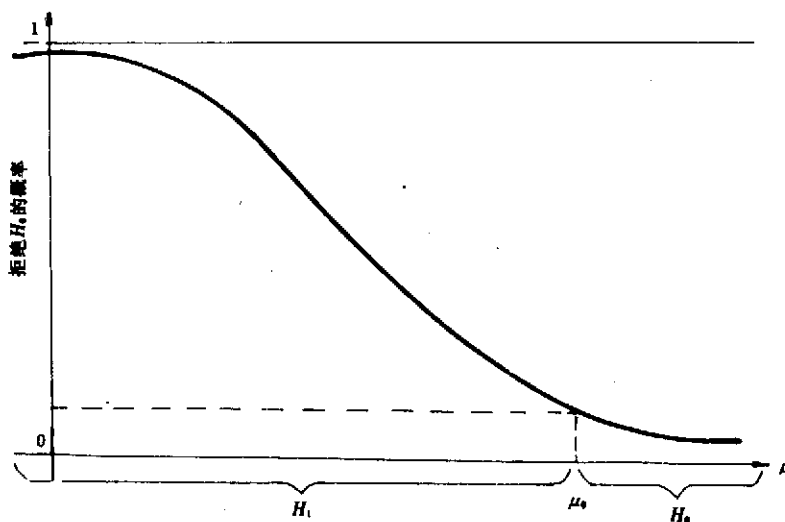


图 1 功效曲线

例 2: 图 2 表示对假设

$$H_0: p < p_0 \leftrightarrow H_1: p \geq p_0$$

的一个检验的功效作为 p 的函数曲线。

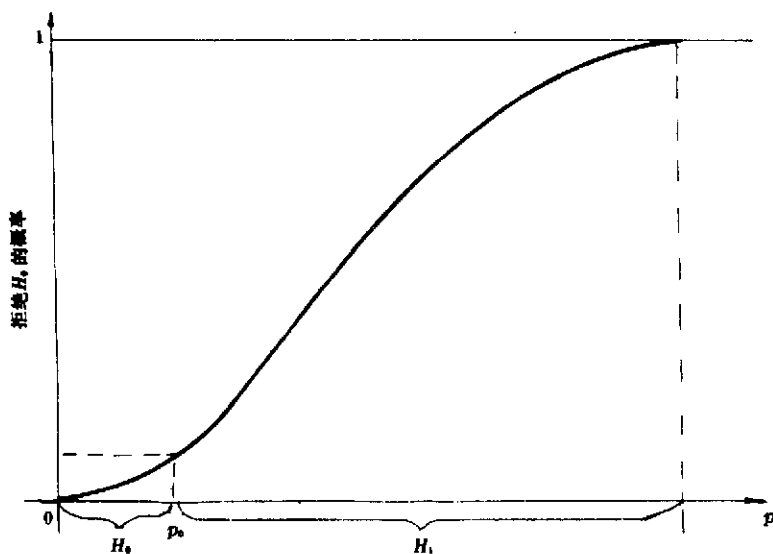


图 2 功效曲线

3.72 特性函数 operating characteristic function

当假设涉及数值参数时,接受原假设的概率是参数的函数,此函数称为检验的特性函数。

同义词: OC 函数 OC function

注: 特性函数等于 1 减功效函数。

3.73 (检验的)特性曲线 operating characteristic curve (of a test)

特性函数的图形表示。

同义词: (检验的)OC 曲线 OC curve (of a test)

例 1: 图 3 是对方差已知,均值 μ 未知的正态分布的假设:

$$H_0: \mu \geq \mu_0 \leftrightarrow H_1: \mu < \mu_0$$

的一个检验的 OC 曲线,它是总体均值 μ 的函数。

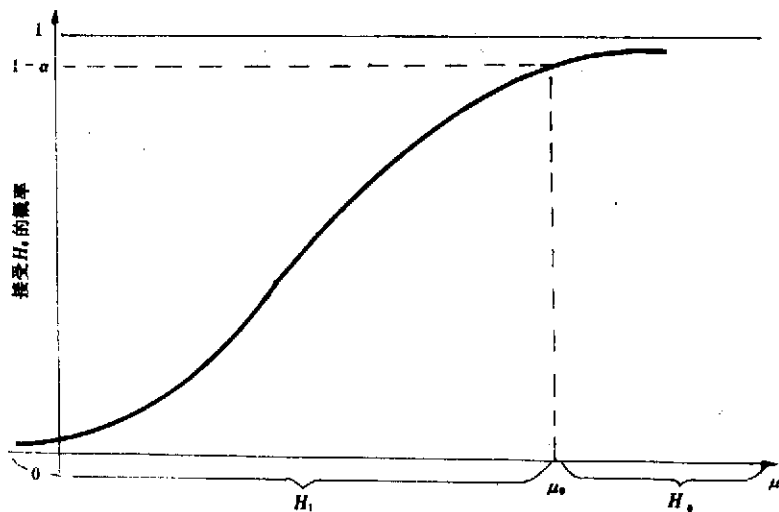


图 3 OC 曲线

例 2: 图 4 是对假设

$$H_0: p < p_0 \leftrightarrow H_1: p > p_0$$

的一个检验的 OC 曲线, 它是 p 的函数。

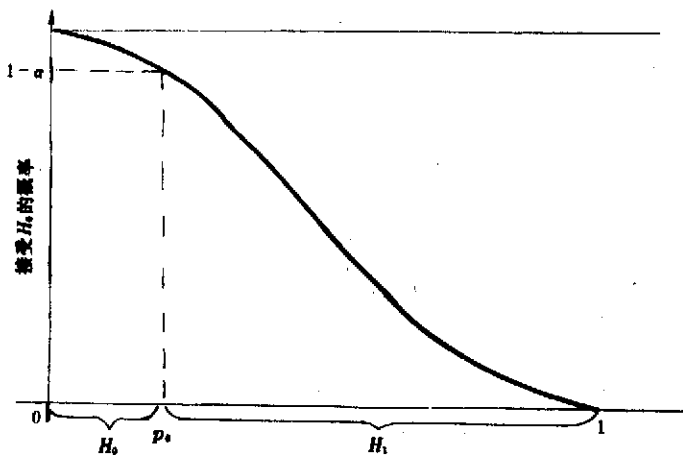


图 4 OC 曲线

3.74 无偏检验 unbiased test

功效不小于显著性水平的检验。

3.75 U 检验 U -test

检验统计量服从正态分布的检验。

例: 对于一个方差已知的正态总体, 关于均值的检验。

3.76 χ^2 检验 chi-square test

检验统计量服从 χ^2 分布的检验。

注: 对某些检验统计量近似服从 χ^2 分布的检验有时也称为 χ^2 检验。

例 1: 通过样本方差检验一个正态总体方差的检验。

例 2: 观测频数与理论频数之间的拟合优度检验。

3.77 t 检验 t -test

检验统计量服从 t 分布的检验。

例 1: 对于一个方差未知的正态总体, 通过样本均值检验总体均值的检验。

例 2: 对于两个方差相等(但未知)的正态总体, 用样本均值差检验总体均值差的检验。

3.78 F 检验 F -test

检验统计量服从 F 分布的检验。

例: 通过样本方差比检验两个正态总体方差比的检验。

3.79 方差分析 analysis of variance

为检验关于模型参数的某些假设或估计方差分量, 把观测值 X_1, X_2, \dots, X_n 总的离差平方和 $\Sigma(X_i - \bar{X})^2$ (其中 \bar{X} 是 X_1, X_2, \dots, X_n 的平均数) 分解为若干个意义明确, 各与特定离差来源有关的部分, 这种方法称为方差分析。

3.80 回归分析 regression analysis

为拟合模型, 通过优化目标函数(例如用最小二乘法)来估计模型的参数, 并对适当的模型假设作检验, 以及用拟合好的模型作统计预报的一种方法。

4 观测和测试结果的一般术语

4.1 [可测]量 [measurable] quantity

可以定性区别并定量确定的用以表述现象、物体或物质的一般属性。

注: 术语“量”可以表示抽象的量, 例如长度, 时间, 质量, 温度, 电阻等; 也可以表示特定的量, 例如一条特定棒的长度, 一段特定导线的电阻等。

4.2 [量的]真值 true value [of a quantity]

一个量在它被观测的瞬时所具有的量值。

注: 一个量的真值仅是一个理想的概念, 通常是未知的。

4.3 [量的]约定真值 conventional true value [of a quantity]

为特定的目的, 用以代替量的真值的量值。

注: ① 约定真值通常被认为是非常接近真值的, 就特定目的而言其差值可以忽略不计。

② 约定真值可用下列方法确定:

- a. 根据理论方法确定;
- b. 由国家或国际性有关组织的试验结果确定;
- c. 由一定科学技术组织主持的合作试验所确认的, 用上述方法确定的约定真值有时也称为参照值或标准值 (reference value)。

4.4 测试结果 test result

用特定的试验方法所确定的量值。

4.5 测试误差 error of a test

测试结果与被测量的真值(或约定真值)之差。

4.6 随机误差 random error

测试误差的组成部分, 在对同一被测量的多次测试中, 它受偶然因素影响而以不可预知的方式变化。

注: 随机误差是不可能被修正的。

4.7 系统误差 systematic error

测试误差的组成部分, 在对同一被测量的多次测量中, 它保持不变或按某种规律而变化。

注: 系统误差及其引起的原因可以是已知的, 也可以是未知的。

4.8 测试的偏倚 bias of a test

测试结果的期望与真值之差。

注：偏倚是由一种或几种系统误差所引起的。

4.9 准确度 accuracy

测试结果与被测量真值或约定真值间的一致程度。

4.10 正确度 trueness

若干次测试结果的平均数的准确程度。

注：正确度表示测试结果中系统误差的大小，通常以偏倚来度量。

4.11 精密度 precision

在规定条件下，相互独立的测试结果之间的一致程度。

注：精密度仅依赖于随机误差，而与被测量的真值或其他约定值无关。

4.12 重复性 repeatability

在重复性条件下(见 4.13 条)，相互独立的测试结果之间的一致程度。

4.13 重复性条件 repeatability conditions

在同一实验室，由同一操作者使用相同设备，按相同的测试程序，并在短时间内对同一被测对象相互独立进行的测试条件。

4.14 重复性标准差 repeatability standard deviation

在重复性条件下，所得测试结果的标准差。

注：可同样定义重复性方差与重复性变异系数。

4.15 重复性限 repeatability limit

一个数值 r ，在重复性条件下，两次测试结果之间的绝对值不超过此数的概率为 95%。

4.16 重复性临界差 repeatability critical difference

一个数值在重复性条件下，两个测试结果或两组测试结果计算结果所得的最后结果(例如平均数，中位数等)之差的绝对值以一个确定的概率不超过此数。

4.17 再现性 reproducibility

在再现性条件(见 4.18 条)下，测试结果之间的一致程度。

同义词：复现性

4.18 再现性条件 reproducibility conditions

在进行测试的实验室，操作者，测试设备，测试程序(方法)，测试时间有所本质变化的情况下，对同一被测对象相互独立进行的测试条件

同义词：复现性条件

注：在确定再现性(复现性)报告中，应当指明再现性条件中变动的情况。

4.19 再现性标准差 reproducibility standard deviation

在再现性条件下，所得测试结果的标准差。

同义词：复现性标准差

注：可同样定义再现性方差与再现性变异系数。

4.20 再现性限 reproducibility limit

一个数值 R ，在再现性条件下，两次测试结果的绝对值不超过此数的概率为 95%。

同义词：复现性限

4.21 再现性临界差 reproducibility critical difference

一个数值，在再现性条件下，两个测试结果或由两组测试结果计算所得的最后结果(例如平均数，中位数等)之差的绝对值以一个确定的概率不超过此数。

同义词：复现性临界差

5 抽样方法的一般术语

5.1 随机抽取 draw an item at random

从由 N 个个体组成的总体中抽取一个个体时,若总体中的每一个体被抽取的可能性都相等,则称这种抽取方法为随机抽取。

注:当总体包含无限多个个体时,“随机抽取”这一概念是通过极限过程由有限总体的情形引伸而来。

5.2 抽样单元 sampling unit

为抽样目的,将总体划分成由个体组成的有限多个部分,每一部分称为一个抽样单元。

注:① 一个抽样单元可以包含一个或多个个体。

② 抽样单元可以分级;总体由初级(抽样)单元组成,每个初级(抽样)单元由二级抽样单元组成,依此类推(见 5.10~5.13 条)。

③ 对散装材料或连续材料进行抽样时,抽样单元定义为具有确定边界的一定量的材料(例如一定容器或包装的材料),或对特定时刻或时间区间内取自流水线上的一定数量的材料。

5.3 抽样框 sampling frame

记录或表明总体所包含抽样单元的名册、清单或地图等。

5.4 放回抽样 sampling with replacement

每个被抽取到并经观测后的抽样单元,在抽取下一个抽样单元前都必须放回总体中去,这种抽样方法称为放回抽样。

注:在这种抽样方法中,同一抽样单元有可能在样本中出现多次。

5.5 不放回抽样 sampling without replacement

从总体中一次同时抽取所需个数的抽样单元,或逐个抽取且已抽取到的单元不放回到总体中去,这种抽样方法称为不放回抽样。

5.6 概率抽样 probability sampling

从包含有 N 个抽样单元的总体中抽取样本量为 n 的样本,若每个可能样本都有一个事先确定的被抽出的概率,则称这种抽样为概率抽样。

5.7 简单随机抽样 simple random sampling

从包含 N 个抽样单元的总体中按不放回抽样抽取 n 个单元,若任何 n 个单元被抽出的概率都相等,也即等于 $1/\binom{N}{n}$,则称这种抽样方法为简单随机抽样。

注:① 简单随机抽样可以用以下的逐个抽取单元的方法进行:第一个样本单元从总体中所有 N 个抽样单元中随机抽取,第二个样本单元从剩下的 $N-1$ 个抽样单元中随机抽取,依此类推。

② 按简单随机抽样得到的样本称为简单随机样本(simple random sample)。

5.8 分层抽样 stratified sampling

将总体分割成互不重叠的子总体(层),在每层中独立地按给定的样本量进行的抽样。

注:在分层抽样中,每层至少需抽取一个抽样单元。

5.9 整群抽样 cluster sampling

将总体分成若干互不重叠的群,每个群由若干个体组成。从总体中抽取若干个群,抽出的群中所有的个体组成样本。这种抽样方法称为整群抽样。

5.10 多级抽样 multistage sampling

一种分级抽取样本的抽样方法。第一级抽样是从总体中抽取初级抽样单元,以后每一级抽样是在上一级抽样单元中抽取次一级的抽样单元。

同义词:多阶抽样

5.11 初级样本 primary sample

在多级抽样的第一级,从总体中抽取的由初级抽样单元组成的样本。

同义词:一级样本

- 5.12 二级样本 secondary sample
在多级抽样的第二级,从初级样本中抽取的由二级抽样单元组成的样本。
- 5.13 末级样本 final sample
在多级抽样的最后一级,从上一级样本中抽取的由最小抽样单元组成的样本。
注:在多级抽样中,样本量通常指末级样本中包含的最小抽样单元(或个体)数。
- 5.14 系统抽样 systematic sampling
将总体中的抽样单元按某种次序排列,在规定的范围内随机抽取一个或一组初始单元,然后按一套规则确定其他样本单元的抽样方法。
- 5.15 等距抽样 periodic systematic sampling
将总体中的 N 个抽样单元按某种次序排列,并编上 1 到 N 的号码,抽取 n 个单元的等距抽样,即是抽取号码为 $h, h+k, h+2k, \dots, h+(n-1)k$ 的 n 个单元,其中 k 是最接近 N/n 的整数, h 是从 1 到 k 的整数中随机抽取的初始单元的号码。
同义词:周期系统抽样
- 5.16 抽样比 sampling fraction
样本量与抽取此样本的总体(或子总体)中所包含的抽样单元总数之比。
- 5.17 散料抽样 bulk sampling
对成批而不易区分抽样单元的散装材料的抽样。
例:对一批煤或水泥的抽样。
- 5.18 子样本 subsample
把一个样本当作总体,从其中抽取的样本。
注:从样本中抽取子样本的方法可以与从总体中抽取该样本的方法相同,也可不同。
- 5.19 样本缩分 sample division
从散料样本中用缩分器等方法取得一个或多个子样本的过程。
- 5.20 缩分样本 divided sample
在散料抽样情形下,用样本缩分方法获得的每个子样本。
- 5.21 交叉子样本 interpenetrating sample
把一个包含 n 个单元的复杂样本,如分层样本、多级样本等,随机地划分为 k 个子样本,每个子样本包含 $m=n/k$ 个单元,由此得到的每个子样本称为交叉子样本。
- 5.22 份样 increment
从较大量的散装材料或连续材料中一次抽取的一定数量的材料。
- 5.23 集样 aggregated sample
从散料中抽取的若干份样的混合样本。
- 5.24 样本准备 sample preparation
将集样通过分取、混合、缩分等程序制成实验室样本或测试样本的过程。
- 5.25 实验室样本 laboratory sample
准备用于实验室检查或检测的样本。
- 5.26 测试样本 test sample
准备用于测试或分析的样本。

附录 A
常用符号
(补充件)

A1 常用符号及其相应术语(见表 A1)

表 A1

符 号	相应术语或含义	参考条目
$P(A), P_r(A)$	事件 A 的概率	2.1
X, Y, \dots	随机变量; 总体中关于某特性的可观测值	2.2, 3.4
x, y, \dots	随机变量的特性值或样本观测值	2.2, 3.9
$F(x), G(x), \dots$	分布函数(在 x 处的值)	2.4
$f(x), g(x), \dots$	连续随机变量的概率密度函数 (在 x 处的值)	2.5
N	总体中包含的个体数	3.2
n	样本量	3.7
R, w	极差	3.27
\bar{X}, \bar{x}	样本均值	3.23
$\mu, E(X)$	随机变量或总体的期望(均值)	2.15
$\sigma^2, V_w(X), V(x)$	随机变量或总体的方差	2.18
σ	随机变量或总体的标准差	2.10
S^2, s^2	样本方差	3.2
S, s	样本标准差	3.30
$Cov(X, Y)$	(X 与 Y 的)协方差	2.26
S_{XY}	(X 与 Y 的)样本协方差	3.32
x_p	X 的 p 分位数	2.12
ρ	相关系数	3.33
$\hat{\theta}$	参数 θ 的估计量	3.40
$N(\mu, \sigma^2)$	均值为 μ , 方差为 σ^2 的正态分布	2.33
$\Phi(x)$	标准正态分布的分布函数	2.34
$\phi(x)$	标准正态分布的密度函数	2.34
χ^2	χ^2 分布(随机变量)	2.35
	注: 如有必要, 可以用下列符号:	
	χ_p^2 : χ^2 分布的 p 分位数	
	$\chi^2(\nu)$: 自由度为 ν 的 χ^2 分布(随机变量)	
	$\chi_p^2(\nu)$: 自由度为 ν 的 χ^2 分布的 p 分位数	

续表 A1

符 号	相应术语或含义	参考条目
t	t 分布(随机变量) 注: 如有必要, 可以用下列符号: t_p : t 分布的 p 分位数 $t(\nu)$: 自由度为 ν 的 t 分布(随机变量) $t_p(\nu)$: 自由度为 ν 的 t 分布的 p 分位数	2.36
F	F 分布(随机变量) 注: 如有必要, 可以用下列符号: F_p : F 分布的 p 分位数 $F(\nu_1, \nu_2)$: 自由度为 ν_1, ν_2 的 F 分布(随机变量) $F_p(\nu_1, \nu_2)$: 自由度为 ν_1, ν_2 的 F 分布的 p 分位数	2.37
α	第一类错误概率; 显著性水平	3.65, 3.68
β	第二类错误概率	3.67
$1-\alpha$	置信水平	3.49
r	重复性限	4.15
R	再现性限(复现性限)	4.20

附 录 B
汉 语 索 引
(补充件)

		独立.....	2.11
	B	独立同分布样本	3.8
备择假设.....	3.55	对数正态分布.....	2.38
边缘分布	2.9	多级抽样.....	5.10
变异系数.....	2.20	多阶抽样.....	5.10
标准差.....	2.19	多项分布.....	2.50
标准化随机变量.....	2.21		
标准误差.....	3.46		
标准正态分布.....	2.34	E	
不放回抽样	5.5	二级样本.....	5.12
		二维正态分布.....	2.49
		二项分布.....	2.45
	C		
测试的偏倚	4.8	F	
测试结果	4.4	方差.....	2.18
测试误差	4.5	方差分析.....	3.79
测试样本.....	5.26	放回抽样	5.4
超几何分布.....	2.48	非参数检验.....	3.59
重复性.....	4.12	F 分布.....	2.37
重复性标准差.....	4.14	t 分布	2.36
重复性临界差.....	4.16	B 分布.....	2.41
重复性条件.....	4.13	Γ 分布.....	2.40
重复性限.....	4.15	χ^2 分布	2.35
抽样	3.6	分布函数	2.4
抽样比.....	5.16	分层抽样	5.8
抽样单元	5.2	分位数.....	2.12
抽样框	5.3	份样.....	5.22
抽样误差.....	3.44	峰度.....	2.29
初级样本.....	5.11	复合假设.....	3.57
次序统计量.....	3.24	复现性.....	4.17
		复现性标准差.....	4.19
	D	复现性临界差.....	4.21
单侧检验.....	3.62	复现性条件.....	4.18
单侧置信区间.....	3.48	复现性限.....	4.20
等距抽样.....	5.15	负二项分布.....	2.46
第二类错误.....	3.66		
第二类错误概率.....	3.67	G	
第一类错误.....	3.64	概率	2.1
第一类错误概率.....	3.65	概率抽样	5.6

概率分布	2.3	累积频数	3.15
概率函数	2.6	离群值	3.52
[概率]密度函数	2.5	离散随机变量	2.6
个体	3.1	联合矩	2.25
功效函数	3.70	连续随机变量	2.5
功效曲线	3.71	量	4.1
估计	3.39	列联表	3.21
估计量	3.40	临界值	3.61
估计量的偏倚	3.42		
估计值	3.41	M	
观测值	3.9	末级样本	5.13
H		N	
OC 函数	3.72	拟合优度	3.51
回归方程	2.31		
回归分析	3.80	P	
回归函数	2.31	偏度	2.30
回归系数	2.31	频率	3.16
J		频数	3.14
极差	3.27	平均绝对差	3.28
集样	5.23	泊松分布	2.47
简单假设	3.56		
简单随机抽样	5.7	Q	
<i>F</i> 检验	3.78	期望	2.15
<i>t</i> 检验	3.77	OC 曲线	3.73
<i>U</i> 检验	3.75		
χ^2 检验	3.76	S	
检验的功效	3.69	散点图	3.20
检验统计量	3.58	散料抽样	5.17
交叉子样本	5.21	实验室样本	5.25
精密度	4.11	双侧检验	3.63
经验分布	3.34	双侧置信区间	3.47
经验回归方程	3.36	随机变量	2.2
经验回归系数	3.37	随机抽取	5.1
矩	2.22	随机误差	4.6
拒绝域	3.60	缩分样本	5.20
绝对矩	2.28		
均方误差	3.43	T	
均匀分布	2.32	特性	3.4
均值	2.15	特性函数	3.72
L		特性曲线	3.73
累积频率	3.17	条件分布	2.10
		条件期望	2.16

条形图.....	3.19	样本准备.....	5.24
统计覆盖区间.....	3.50	一级样本.....	5.11
统计假设.....	3.53	游程.....	3.38
统计检验.....	3.54	原点矩.....	2.23
统计量.....	3.22	原假设.....	3.55
		约定真值	4.3
W		Z	
威布尔分布.....	2.44	再现性.....	4.17
k 维概率分布	2.8	再现性标准差.....	4.19
k 维随机变量	2.7	再现性临界差.....	4.21
无偏估计量.....	3.45	再现性条件.....	4.18
无偏检验.....	3.74	再现性限.....	4.20
		真值	4.2
X		整群抽样	5.9
系统抽样.....	5.14	正确度.....	4.10
系统误差	4.7	正态分布.....	2.33
显著性水平.....	3.68	直方图.....	3.18
相关系数.....	2.27	指数分布.....	2.39
协方差.....	2.26	置信水平.....	3.49
I 型极值分布.....	2.42	中程数.....	3.26
II 型极值分布.....	2.43	中位数.....	2.13
		中心化随机变量.....	2.17
Y		中心矩.....	2.24
样本	3.5	众数.....	2.14
样本变异系数.....	3.31	周期系统抽样.....	5.15
样本标准差.....	3.30	准确度	4.9
样本方差.....	3.29	子样本.....	5.18
样本矩.....	3.35	总体	3.2
样本均值.....	3.23	总体分布	3.3
样本量	3.7	组.....	3.10
样本缩分.....	5.19	组矩.....	3.13
样本相关系数.....	3.33	组限.....	3.11
样本协方差.....	3.32	组中值.....	3.12
样本中位数.....	3.25		

附 录 C
英 文 索 引
(补充件)

A

absolute frequency	3.14
absolute moment	2.28
accuracy	4.9
aggregated sample	5.23
alternative hypothesis	3.55
analysis of variance	3.79

B

bar chart	3.19
beta distribution	2.41
bias of a test	4.8
bias of estimator	3.42
binomial distribution	2.45
bivariate normal distribution	2.49
bulk sampling	5.17

C

central moment	2.24
central random variable	2.17
characteristic	3.4
chi-square distribution	2.35
chi-square test	3.76
class	3.10
class limits	3.11
class width	3.13
cluster sampling	5.9
coefficient of variation	2.20
composite hypothesis	3.57
conditional distribution	2.10
conditional expectation	2.16
confidence level	3.49
contingency table	3.21
continuous random variable	2.5
conventional true value	4.3
correlation coefficient	2.27
covariance	2.26

critical value	3.61
cumulative absolute frequency	3.15
cumulative relative frequency	3.17

D

density function	2.5
discrete random variable	2.6
distribution function	2.4
divided sample	5.20
draw an item at random	5.1

E

empirical distribution	3.34
empirical regression coefficient	3.37
empirical regression equation	3.36
error of a test	4.5
estimate	3.41
estimation	3.39
estimator	3.40
expectation	2.15
exponential distribution	2.39

F

F-distribution	2.37
final sample	5.13
Fréchet distribution	2.43
F-test	3.78

G

gamma distribution	2.40
goodness of fit	3.51
Gumbel distribution	2.42

H

histogram	3.18
hypergeometric distribution	2.48

I

increment	5.22
independence	2.11
independently identically distributed sample	3.8
individual	3.1
interpenetrating sample	5.21

item	3.1
J	
joint moment	2.25
K	
k-dimensional probability distribution	2.8
k-dimensional random variable	2.7
kurtosis	2.29
L	
laboratory sample	5.25
log-normal distribution	2.38
M	
marginal distribution	2.9
mean	2.15
mean deviation	3.28
mean square error	3.43
median	2.13
mid-point of class	3.12
midrange	3.26
mode	2.14
moment	2.22
moment about the origin	2.23
multinomial distribution	2.50
multistage sampling	5.10
N	
negative binomial distribution	2.46
non-parametric test	3.59
normal distribution	2.33
null hypothesis	3.55
O	
observed value	3.9
OC curve	3.73
OC function	3.72
one-sided confidence interval	3.48
one-sided test	3.62
operating characteristic curve	3.73
operating characteristic function	3.72
order statistics	3.24

outlier	3.52
P	
periodic systematic sampling	5.15
Poisson distribution	2.47
population	3.2
population distribution	3.3
power curve	3.71
power function	3.70
power of a test	3.69
precision	4.11
primary sample	5.11
probability	2.1
probability distribution	2.3
probability function	2.6
probability sampling	5.6
Q	
quantile	2.12
quantity	4.1
R	
random error	4.6
random variable	2.2
range	3.27
regression analysis	3.80
regression coefficient	2.31
regression equation	2.31
regression function	2.31
rejection region	3.60
relative frequency	3.16
repeatability	4.12
repeatability conditions	4.13
repeatability critical difference	4.16
repeatability limit	4.15
repeatability standard deviation	4.14
reproducibility	4.17
reproducibility conditions	4.18
reproducibility critical difference	4.21
reproducibility limit	4.20
reproducibility standard deviation	4.19
run	3.38

S

sample	3.5
sample coefficient of variation	3.31
sample correlation coefficient	3.33
sample covariance	3.32
sample division	5.19
sample mean	3.23
sample median	3.25
sample moment	3.35
sample preparation	5.24
sample size	3.7
sample standard deviation	3.30
sample variance	3.29
sampling	3.6
sampling error	3.44
sampling fraction	5.16
sampling frame	5.3
sampling unit	5.2
sampling with replacement	5.4
sampling without replacement	5.5
scatter diagram	3.20
secondary sample	5.12
significant level	3.68
simple hypothesis	3.56
simple random sampling	5.7
skewness	2.30
standard deviation	2.19
standard error	3.46
standardized normal distribution	2.34
standardized random variable	2.21
statistic	3.22
statistical coverage interval	3.50
statistical hypothesis	3.53
statistical test	3.54
stratified sampling	5.8
Student's distribution	2.36
subsample	5.18
systematic error	4.7
systematic sampling	5.14

T

t-distribution	2.36
----------------------	------

test result	4.4
test sample	5.26
test statistic	3.58
true value	4.2
trueness	4.10
t-test	3.77
two-sided confidence interval	3.47
two-sided test	3.63
type I error	3.64
type I error probability	3.65
type I extreme value distribution	2.42
type II error	3.66
type II error probability	3.67
type II extreme value distribution	2.43
type III extreme value distribution	2.44

U

unbiased estimator	3.45
unbiased test	3.74
uniform distribution	2.32
<i>U</i> -test	3.75

V

variance	2.18
variate	2.2

W

Weibull distribution	2.44
----------------------------	------

附加说明：

本标准由全国统计方法应用标准化技术委员会提出。

本标准由全国统计方法应用标准化技术委员会术语、符号和统计用表分委员会工作组负责起草。

本标准主要起草人陈希孺、陶波、冯士雍、吴启光。