

第 1 章 随机事件及其概率

随 机 试 验 和 随 机 事 件	<p>如果一个试验在相同条件下可以重复进行，而每次试验的可能结果不止一个，但在进行一次试验之前却不能断言它出现哪个结果，则称这种试验为随机试验。</p> <p>试验的可能结果称为随机事件。</p>
基本事件、 样 本 空 间 和事件	<p>在一个试验下，不管事件有多少个，总可以从其中找出这样一组事件，它具有如下性质：</p> <p>①每进行一次试验，必须发生且只能发生这一组中的一个事件；</p> <p>②任何事件，都是由这一组中的部分事件组成的。</p> <p>这样一组事件中的每一个事件称为基本事件，用 ω 来表示。</p> <p>基本事件的全体，称为试验的样本空间，用 Ω 表示。</p> <p>一个事件就是由 Ω 中的部分点（基本事件 ω）组成的集合。通常用大写字母 A, B, C, \dots 表示事件，它们是 Ω 的子集。</p> <p>Ω 为必然事件，\emptyset 为不可能事件。</p> <p>不可能事件 (\emptyset) 的概率为零，而概率为零的事件不一定是不可事件；同理，必然事件 (Ω) 的概率为 1，而概率为 1 的事件也不一定是必然事件。</p>
) 事件的关 系与运算	<p>①关系：</p> <p>如果事件 A 的组成部分也是事件 B 的组成部分，(A 发生必有事件 B 发生)：</p> $A \subset B$ <p>如果同时有 $A \subset B, B \supset A$，则称事件 A 与事件 B 等价，或称 A 等于 B：$A=B$。</p> <p>A, B 中至少有一个发生的事件：$A \cup B$，或者 $A+B$。</p> <p>属于 A 而不属于 B 的部分所构成的事件，称为 A 与 B 的差，记为 $A-B$，也可表示为 $A-AB$ 或者 \overline{AB}，它表示 A 发生而 B 不发生的事件。</p> <p>A, B 同时发生：$A \cap B$，或者 AB。$A \cap B = \emptyset$，则表示 A 与 B 不可能同时发生，称事件 A 与事件 B 互不相容或者互斥。基本事件是互不相容的。</p> <p>$\Omega - A$ 称为事件 A 的逆事件，或称 A 的对立事件，记为 \overline{A}。它表示 A 不发生的事件。互斥未必对立。</p> <p>②运算：</p> <p>结合率：$A(BC)=(AB)C \quad A \cup (B \cup C)=(A \cup B) \cup C$</p> <p>分配率：$(AB) \cup C=(A \cup C) \cap (B \cup C) \quad (A \cup B) \cap C=(AC) \cup (BC)$</p> $\overline{\bigcap_{i=1}^{\infty} A_i} = \bigcup_{i=1}^{\infty} \overline{A_i}$ <p>德摩根率：$\overline{A \cup B} = \overline{A} \cap \overline{B}, \quad \overline{A \cap B} = \overline{A} \cup \overline{B}$</p>
概 率 的 公 理 化 定 义	<p>设 Ω 为样本空间，A 为事件，对每一个事件 A 都有一个实数 $P(A)$，若满足下列三个条件：</p> <p>1° $0 \leq P(A) \leq 1$,</p> <p>2° $P(\Omega) = 1$</p> <p>3° 对于两两互不相容的事件 A_1, A_2, \dots 有</p>

	$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$ <p>常称为可列（完全）可加性。</p> <p>则称 $P(A)$ 为事件 A 的概率。</p>
古典概型	<p>1° $\Omega = \{\omega_1, \omega_2 \cdots \omega_n\}$,</p> <p>2° $P(\omega_1) = P(\omega_2) = \cdots P(\omega_n) = \frac{1}{n}$。</p> <p>设任一事件 A，它是由 $\omega_1, \omega_2 \cdots \omega_m$ 组成的，则有</p> $P(A) = \{(\omega_1) \cup (\omega_2) \cup \cdots \cup (\omega_m)\} = P(\omega_1) + P(\omega_2) + \cdots + P(\omega_m)$ $= \frac{m}{n} = \frac{A \text{ 所包含的基本事件数}}{\text{基本事件总数}}$
几何概型	<p>若随机试验的结果为无限不可数并且每个结果出现的可能性均匀，同时样本空间中的每一个基本事件可以使用一个有界区域来描述，则称此随机试验为几何概型。对任一事件 A,</p> $P(A) = \frac{L(A)}{L(\Omega)}$ <p>其中 L 为几何度量（长度、面积、体积）。</p>
加法公式	$P(A+B) = P(A) + P(B) - P(AB)$ <p>当 $P(AB) = 0$ 时，$P(A+B) = P(A) + P(B)$</p>
减法公式	$P(A-B) = P(A) - P(AB)$ <p>当 $B \subset A$ 时，$P(A-B) = P(A) - P(B)$</p> <p>当 $A = \Omega$ 时，$P(\bar{B}) = 1 - P(B)$</p>
条件概率	<p>定义 设 A、B 是两个事件，且 $P(A) > 0$，则称 $\frac{P(AB)}{P(A)}$ 为事件 A 发生条件下，事件 B 发生的条件概率，记为 $P(B/A) = \frac{P(AB)}{P(A)}$。</p> <p>条件概率是概率的一种，所有概率的性质都适合于条件概率。</p> <p>例如 $P(\Omega/B) = 1 \Rightarrow P(\bar{B}/A) = 1 - P(B/A)$</p>
乘法公式	<p>乘法公式：$P(AB) = P(A)P(B/A)$</p> <p>更一般地，对事件 $A_1, A_2, \cdots A_n$，若 $P(A_1 A_2 \cdots A_{n-1}) > 0$，则有</p> $P(A_1 A_2 \cdots A_n) = P(A_1)P(A_2 A_1)P(A_3 A_1 A_2) \cdots P(A_n A_1 A_2 \cdots A_{n-1})$
独立性	<p>①两个事件的独立性</p> <p>设事件 A、B 满足 $P(AB) = P(A)P(B)$，则称事件 A、B 是相互独立的。</p> <p>若事件 A、B 相互独立，且 $P(A) > 0$，则有</p> $P(B A) = \frac{P(AB)}{P(A)} = \frac{P(A)P(B)}{P(A)} = P(B)$ <p>若事件 A、B 相互独立，则可得到 \bar{A} 与 B、A 与 \bar{B}、\bar{A} 与 \bar{B} 也都相互独立。</p>

	<p>必然事件 Ω 和不可能事件 \emptyset 与任何事件都相互独立。 \emptyset 与任何事件都互斥。</p> <p>②多个事件的独立性 设 ABC 是三个事件, 如果满足两两独立的条件, $P(AB)=P(A)P(B)$; $P(BC)=P(B)P(C)$; $P(CA)=P(C)P(A)$ 并且同时满足 $P(ABC)=P(A)P(B)P(C)$ 那么 A、B、C 相互独立。 对于 n 个事件类似。</p>
全概公式	<p>设事件 B_1, B_2, \dots, B_n 满足</p> <p>1° B_1, B_2, \dots, B_n 两两互不相容, $P(B_i) > 0 (i = 1, 2, \dots, n)$,</p> $A \subset \bigcup_{i=1}^n B_i$ <p>2° ,</p> <p>则有</p> $P(A) = P(B_1)P(A B_1) + P(B_2)P(A B_2) + \dots + P(B_n)P(A B_n)。$
贝叶斯公式	<p>设事件 B_1, B_2, \dots, B_n 及 A 满足</p> <p>1° B_1, B_2, \dots, B_n 两两互不相容, $P(B_i) > 0, i = 1, 2, \dots, n$,</p> $A \subset \bigcup_{i=1}^n B_i$ <p>2° , $P(A) > 0$,</p> <p>则</p> $P(B_i A) = \frac{P(B_i)P(A B_i)}{\sum_{j=1}^n P(B_j)P(A B_j)}, i=1, 2, \dots, n。$ <p>此公式即为贝叶斯公式。</p> <p>$P(B_i), (i = 1, 2, \dots, n)$, 通常叫先验概率。$P(B_i A), (i = 1, 2, \dots, n)$, 通常称为后验概率。贝叶斯公式反映了“因果”的概率规律, 并作出了“由果溯因”的推断。</p>
伯努利概型	<p>我们作了 n 次试验, 且满足</p> <ul style="list-style-type: none"> ◆ 每次试验只有两种可能结果, A 发生或 A 不发生; ◆ n 次试验是重复进行的, 即 A 发生的概率每次均一样; ◆ 每次试验是独立的, 即每次试验 A 发生与否与其他次试验 A 发生与否是互不影响的。 <p>这种试验称为伯努利概型, 或称为 n 重伯努利试验。</p> <p>用 p 表示每次试验 A 发生的概率, 则 \bar{A} 发生的概率为 $1-p=q$, 用 $P_n(k)$ 表示 n 重伯努利试验中 A 出现 $k(0 \leq k \leq n)$ 次的概率,</p> $P_n(k) = C_n^k p^k q^{n-k}, k = 0, 1, 2, \dots, n。$

第二章 随机变量及其分布

(1) 离散型随机变量的分布律	<p>设离散型随机变量 X 的可能取值为 $X_k(k=1,2,\cdots)$ 且取各个值的概率, 即事件 $(X=X_k)$ 的概率为</p> $P(X=x_k)=p_k, \quad k=1,2,\cdots,$ <p>则称上式为离散型随机变量 X 的概率分布或分布律。有时也用分布列的形式给出:</p> $\begin{array}{c c} X & x_1, x_2, \cdots, x_k, \cdots \\ \hline P(X=x_k) & p_1, p_2, \cdots, p_k, \cdots \end{array}.$ <p>显然分布律应满足下列条件:</p> $(1) \quad p_k \geq 0, \quad k=1,2,\cdots, \quad (2) \quad \sum_{k=1}^{\infty} p_k = 1.$
(2) 连续型随机变量的分布密度	<p>设 $F(x)$ 是随机变量 X 的分布函数, 若存在非负函数 $f(x)$, 对任意实数 x, 有</p> $F(x) = \int_{-\infty}^x f(x)dx,$ <p>则称 X 为连续型随机变量。 $f(x)$ 称为 X 的概率密度函数或密度函数, 简称概率密度。</p> <p>密度函数具有下面 4 个性质:</p> $1. \quad f(x) \geq 0.$ $2. \quad \int_{-\infty}^{+\infty} f(x)dx = 1.$
(3) 离散与连续型随机变量的关系	$P(X=x) \approx P(x < X \leq x+dx) \approx f(x)dx$ <p>积分元 $f(x)dx$ 在连续型随机变量理论中所起的作用与 $P(X=x_k)=p_k$ 在离散型随机变量理论中所起的作用相类似。</p>

(4) 分布函数	<p>设 X 为随机变量, x 是任意实数, 则函数</p> $F(x) = P(X \leq x)$ <p>称为随机变量 X 的分布函数, 本质上是一个累积函数。</p> $P(a < X \leq b) = F(b) - F(a)$ 可以得到 X 落入区间 $(a, b]$ 的概率。分布函数 $F(x)$ 表示随机变量落入区间 $(-\infty, x]$ 内的概率。 <p>分布函数具有如下性质:</p> <p>1° $0 \leq F(x) \leq 1, \quad -\infty < x < +\infty;$</p> <p>2° $F(x)$ 是单调不减的函数, 即 $x_1 < x_2$ 时, 有 $F(x_1) \leq F(x_2);$</p> <p>3° $F(-\infty) = \lim_{x \rightarrow -\infty} F(x) = 0, \quad F(+\infty) = \lim_{x \rightarrow +\infty} F(x) = 1;$</p> <p>4° $F(x+0) = F(x)$, 即 $F(x)$ 是右连续的;</p> <p>5° $P(X = x) = F(x) - F(x-0)。$</p> <p>对于离散型随机变量, $F(x) = \sum_{x_k \leq x} p_k;$</p> <p>对于连续型随机变量, $F(x) = \int_{-\infty}^x f(x)dx。$</p>	
(5) 八大分布	0-1 分布	$P(X=1)=p, P(X=0)=q$
	二项分布	<p>在 n 重贝努里试验中, 设事件 A 发生的概率为 p。事件 A 发生的次数是随机变量, 设为 X, 则 X 可能取值为 $0, 1, 2, \dots, n$。</p> $P(X = k) = P_n(k) = C_n^k p^k q^{n-k}, \quad \text{其中}$ $q = 1 - p, 0 < p < 1, k = 0, 1, 2, \dots, n,$ <p>则称随机变量 X 服从参数为 n, p 的二项分布。记为 $X \sim B(n, p)。$</p> <p>当 $n = 1$ 时, $P(X = k) = p^k q^{1-k}, k = 0, 1$, 这就是 (0-1) 分布, 所以 (0-1) 分布是二项分布的特例。</p>

泊松分布	<p>设随机变量 X 的分布律为</p> $P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad \lambda > 0, \quad k = 0, 1, 2, \dots,$ <p>则称随机变量 X 服从参数为 λ 的泊松分布, 记为 $X \sim \pi(\lambda)$ 或者 $P(\lambda)$。</p> <p>泊松分布为二项分布的极限分布 ($np = \lambda, n \rightarrow \infty$)。</p>
超几何分布	$P(X = k) = \frac{C_M^k \cdot C_{N-M}^{n-k}}{C_N^n}, \quad k = 0, 1, 2, \dots, l$ <p>随机变量 X 服从参数为 n, N, M 的超几何分布, 记为 $H(n, N, M)$。</p>
几何分布	$P(X = k) = q^{k-1} p, \quad k = 1, 2, 3, \dots,$ <p>其中 $p \geq 0, q = 1 - p$。</p> <p>随机变量 X 服从参数为 p 的几何分布, 记为 $G(p)$。</p>
均匀分布	<p>设随机变量 X 的值只落在 $[a, b]$ 内, 其密度函数 $f(x)$ 在 $[a, b]$ 上为常数 $\frac{1}{b-a}$, 即</p> $f(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{其他,} \end{cases}$ <p>则称随机变量 X 在 $[a, b]$ 上服从均匀分布, 记为 $X \sim U(a, b)$。</p> <p>分布函数为</p> $F(x) = \int_{-\infty}^x f(x) dx = \begin{cases} 0, & x < a, \\ \frac{x-a}{b-a}, & a \leq x \leq b \\ 1, & x > b. \end{cases}$ <p>当 $a \leq x_1 < x_2 \leq b$ 时, X 落在区间 (x_1, x_2) 内的概率为</p> $P(x_1 < X < x_2) = \frac{x_2 - x_1}{b - a}.$

	指数分布	$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0, \\ 0, & x < 0, \end{cases}$ <p>其中 $\lambda > 0$，则称随机变量 X 服从参数为 λ 的指数分布。 X 的分布函数为</p> $F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0, \\ 0, & x < 0. \end{cases}$ <p>记住积分公式:</p> $\int_0^{+\infty} x^n e^{-x} dx = n!$
--	------	---

	正态分布	<p>设随机变量 X 的密度函数为</p> $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < +\infty,$ <p>其中 μ、$\sigma > 0$ 为常数, 则称随机变量 X 服从参数为 μ、σ 的正态分布或高斯 (Gauss) 分布, 记为 $X \sim N(\mu, \sigma^2)$。</p> <p>$f(x)$ 具有如下性质:</p> <p>1° $f(x)$ 的图形是关于 $x = \mu$ 对称的;</p> <p>2° 当 $x = \mu$ 时, $f(\mu) = \frac{1}{\sqrt{2\pi}\sigma}$ 为最大值;</p> <p>若 $X \sim N(\mu, \sigma^2)$, 则 X 的分布函数为</p> $F(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x e^{-\frac{t^2}{2\sigma^2}} dt$ <p>参数 $\mu = 0$、$\sigma = 1$ 时的正态分布称为标准正态分布, 记为 $X \sim N(0,1)$, 其密度函数记为</p> $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad -\infty < x < +\infty,$ <p>分布函数为</p> $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$ <p>$\Phi(x)$ 是不可求积函数, 其函数值, 已编制成表可供查用。</p> <p>$\Phi(-x) = 1 - \Phi(x)$ 且 $\Phi(0) = \frac{1}{2}$。</p> <p>如果 $X \sim N(\mu, \sigma^2)$, 则 $\frac{X - \mu}{\sigma} \sim N(0,1)$。</p> $P(x_1 < X \leq x_2) = \Phi\left(\frac{x_2 - \mu}{\sigma}\right) - \Phi\left(\frac{x_1 - \mu}{\sigma}\right)。$								
(6) 分位数	下分位表: $P(X \leq \mu_a) = a$;									
(7) 函数分布	离散型	<p>已知 X 的分布列为</p> <table><tr><td>X</td><td>$x_1, x_2, \cdots, x_n, \cdots$</td></tr><tr><td>$P(X = x_i)$</td><td>$p_1, p_2, \cdots, p_n, \cdots$</td></tr></table> <p>$Y = g(X)$ 的分布列 ($y_i = g(x_i)$ 互不相等) 如下:</p> <table><tr><td>Y</td><td>$g(x_1), g(x_2), \cdots, g(x_n), \cdots$</td></tr><tr><td>$P(Y = y_i)$</td><td>$p_1, p_2, \cdots, p_n, \cdots$</td></tr></table> <p>若有某些 $g(x_i)$ 相等, 则应将对应的 p_i 相加作为 $g(x_i)$ 的概率。</p>	X	$x_1, x_2, \cdots, x_n, \cdots$	$P(X = x_i)$	$p_1, p_2, \cdots, p_n, \cdots$	Y	$g(x_1), g(x_2), \cdots, g(x_n), \cdots$	$P(Y = y_i)$	$p_1, p_2, \cdots, p_n, \cdots$
X	$x_1, x_2, \cdots, x_n, \cdots$									
$P(X = x_i)$	$p_1, p_2, \cdots, p_n, \cdots$									
Y	$g(x_1), g(x_2), \cdots, g(x_n), \cdots$									
$P(Y = y_i)$	$p_1, p_2, \cdots, p_n, \cdots$									
	连续型	<p>先利用 X 的概率密度 $f_X(x)$ 写出 Y 的分布函数 $F_Y(y) = P(g(X) \leq y)$, 再利用变上下限积分的求导公式求出 $f_Y(y)$。</p>								

第三章 二维随机变量及其分布

(1) 联合分布	离散型	<p>如果二维随机向量 ξ (X, Y) 的所有可能取值为至多可列个有序对 (x,y), 则称 ξ 为离散型随机量。</p> <p>设 $\xi = (X, Y)$ 的所有可能取值为 $(x_i, y_j)(i, j = 1, 2, \cdots)$, 且事件 $\{\xi = (x_i, y_j)\}$ 的概率为 p_{ij}, 称</p> $P\{(X, Y) = (x_i, y_j)\} = p_{ij} (i, j = 1, 2, \cdots)$ <p>为 $\xi = (X, Y)$ 的分布律或称为 X 和 Y 的联合分布律。联合分布有时也用下面的概率分布表来表示:</p> <table><tr><th>$\begin{matrix} Y \\ X \end{matrix}$</th><th>$y_1$</th><th>$y_2$</th><th>$\cdots$</th><th>$y_j$</th><th>$\cdots$</th></tr><tr><th>$x_1$</th><td>$p_{11}$</td><td>$p_{12}$</td><td>$\cdots$</td><td>$p_{1j}$</td><td>$\cdots$</td></tr><tr><th>$x_2$</th><td>$p_{21}$</td><td>$p_{22}$</td><td>$\cdots$</td><td>$p_{2j}$</td><td>$\cdots$</td></tr><tr><th>$\vdots$</th><td>$\vdots$</td><td>$\vdots$</td><td></td><td>$\vdots$</td><td>$\vdots$</td></tr><tr><th>$x_i$</th><td>$p_{i1}$</td><td></td><td>$\cdots$</td><td>$p_{ij}$</td><td>$\cdots$</td></tr><tr><th>$\vdots$</th><td>$\vdots$</td><td>$\vdots$</td><td></td><td>$\vdots$</td><td>$\vdots$</td></tr></table> <p>这里 p_{ij} 具有下面两个性质:</p> <p>(1) $p_{ij} \geq 0$ (i,j=1,2,\cdots);</p> <p>(2) $\sum_i \sum_j p_{ij} = 1$.</p>	$\begin{matrix} Y \\ X \end{matrix}$	y_1	y_2	\cdots	y_j	\cdots	x_1	p_{11}	p_{12}	\cdots	p_{1j}	\cdots	x_2	p_{21}	p_{22}	\cdots	p_{2j}	\cdots	\vdots	\vdots	\vdots		\vdots	\vdots	x_i	p_{i1}		\cdots	p_{ij}	\cdots	\vdots	\vdots	\vdots		\vdots	\vdots
$\begin{matrix} Y \\ X \end{matrix}$	y_1	y_2	\cdots	y_j	\cdots																																	
x_1	p_{11}	p_{12}	\cdots	p_{1j}	\cdots																																	
x_2	p_{21}	p_{22}	\cdots	p_{2j}	\cdots																																	
\vdots	\vdots	\vdots		\vdots	\vdots																																	
x_i	p_{i1}		\cdots	p_{ij}	\cdots																																	
\vdots	\vdots	\vdots		\vdots	\vdots																																	
	连续型	<p>对于二维随机向量 $\xi = (X, Y)$, 如果存在非负函数 $f(x, y)(-\infty < x < +\infty, -\infty < y < +\infty)$, 使对任意一个其邻边分别平行于坐标轴的矩形区域 D, 即 $D = \{(X, Y) a < x < b, c < y < d\}$ 有</p> $P\{(X, Y) \in D\} = \iint_D f(x, y) dx dy,$ <p>则称 ξ 为连续型随机向量; 并称 $f(x, y)$ 为 $\xi = (X, Y)$ 的分布密度或称为 X 和 Y 的联合分布密度。</p> <p>分布密度 $f(x, y)$ 具有下面两个性质:</p> <p>(1) $f(x, y) \geq 0$;</p> <p>(2) $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = 1$.</p>																																				
(2) 二维随机变量的本质	$\xi(X = x, Y = y) = \xi(X = x \cap Y = y)$																																					

(3) 联合分布函数	<p>设 (X, Y) 为二维随机变量, 对于任意实数 x, y, 二元函数</p> $F(x, y) = P\{X \leq x, Y \leq y\}$ <p>称为二维随机向量 (X, Y) 的分布函数, 或称为随机变量 X 和 Y 的联合分布函数。</p> <p>分布函数是一个以全平面为其定义域, 以事件 $\{(\omega_1, \omega_2) -\infty < X(\omega_1) \leq x, -\infty < Y(\omega_2) \leq y\}$ 的概率为函数值的一个实值函数。分布函数 $F(x, y)$ 具有以下的基本性质:</p> <p>(1) $0 \leq F(x, y) \leq 1$;</p> <p>(2) $F(x, y)$ 分别对 x 和 y 是非减的, 即当 $x_2 > x_1$ 时, 有 $F(x_2, y) \geq F(x_1, y)$; 当 $y_2 > y_1$ 时, 有 $F(x, y_2) \geq F(x, y_1)$;</p> <p>(3) $F(x, y)$ 分别对 x 和 y 是右连续的, 即</p> $F(x, y) = F(x + 0, y), F(x, y) = F(x, y + 0);$ <p>(4) $F(-\infty, -\infty) = F(-\infty, y) = F(x, -\infty) = 0, F(+\infty, +\infty) = 1$.</p> <p>(5) 对于 $x_1 < x_2, y_1 < y_2$,</p> $F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x_1, y_1) \geq 0.$	
(4) 离散型与连续型的关系	$P(X = x, Y = y) \approx P(x < X \leq x + dx, y < Y \leq y + dy) \approx f(x, y) dx dy$	
(5) 边缘分布	离散型	<p>X 的边缘分布为</p> $P_{i\bullet} = P(X = x_i) = \sum_j p_{ij} (i, j = 1, 2, \dots);$ <p>Y 的边缘分布为</p> $P_{\bullet j} = P(Y = y_j) = \sum_i p_{ij} (i, j = 1, 2, \dots).$
	连续型	<p>X 的边缘分布密度为</p> $f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy;$ <p>Y 的边缘分布密度为</p> $f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx.$

(6) 条件分布	离散型	在已知 $X=x_i$ 的条件下, Y 取值的条件分布为 $P(Y = y_j X = x_i) = \frac{p_{ij}}{p_{i\bullet}};$ 在已知 $Y=y_j$ 的条件下, X 取值的条件分布为 $P(X = x_i Y = y_j) = \frac{p_{ij}}{p_{\bullet j}},$
	连续型	在已知 $Y=y$ 的条件下, X 的条件分布密度为 $f(x y) = \frac{f(x,y)}{f_Y(y)};$ 在已知 $X=x$ 的条件下, Y 的条件分布密度为 $f(y x) = \frac{f(x,y)}{f_X(x)}$
(7) 独立性	一般型	$F(X,Y)=F_X(x)F_Y(y)$
	离散型	$p_{ij} = p_{i\bullet} p_{\bullet j}$ 有零不独立
	连续型	$f(x,y)=f_X(x)f_Y(y)$ 直接判断, 充要条件: ①可分离变量 ②正概率密度区间为矩形
	二维正态分布	$f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]},$ $\rho = 0$
	随机变量的函数	若 $X_1, X_2, \dots, X_m, X_{m+1}, \dots, X_n$ 相互独立, h,g 为连续函数, 则: $h(X_1, X_2, \dots, X_m)$ 和 $g(X_{m+1}, \dots, X_n)$ 相互独立。 特例: 若 X 与 Y 独立, 则: h(X) 和 g(Y) 独立。 例如: 若 X 与 Y 独立, 则: $3X+1$ 和 $5Y-2$ 独立。

(8) 二维
均匀分布

设随机向量 (X, Y) 的分布密度函数为

$$f(x, y) = \begin{cases} \frac{1}{S_D} & (x, y) \in D \\ 0, & \text{其他} \end{cases}$$

其中 S_D 为区域 D 的面积, 则称 (X, Y) 服从 D 上的均匀分布, 记为 $(X, Y) \sim U(D)$ 。

例如图 3.1、图 3.2 和图 3.3。

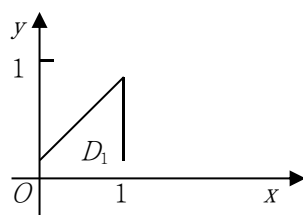


图 3.1

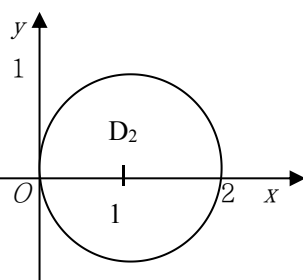


图 3.2

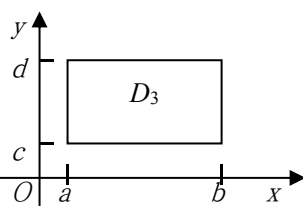


图 3.3

<p>(9) 二维正态分布</p>	<p>设随机向量 (X, Y) 的分布密度函数为</p> $f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]},$ <p>其中 $\mu_1, \mu_2, \sigma_1 > 0, \sigma_2 > 0, \rho < 1$ 是 5 个参数, 则称 (X, Y) 服从二维正态分布,</p> <p>记为 $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$.</p> <p>由边缘密度的计算公式, 可以推出二维正态分布的两个边缘分布仍为正态分布,</p> <p>即 $X \sim N(\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2)$.</p> <p>但是若 $X \sim N(\mu_1, \sigma_1^2), Y \sim N(\mu_2, \sigma_2^2)$, (X, Y) 未必是二维正态分布。</p>	
<p>(10) 函数分布</p>	<p>$Z=X+Y$</p>	<p>根据定义计算: $F_Z(z) = P(Z \leq z) = P(X + Y \leq z)$</p> <p>对于连续型, $f_Z(z) = \int_{-\infty}^{+\infty} f(x, z-x)dx$</p> <p>两个独立的正态分布的和仍为正态分布 $(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$。</p> <p>$n$ 个相互独立的正态分布的线性组合, 仍服从正态分布。</p> $\mu = \sum_i C_i \mu_i, \quad \sigma^2 = \sum_i C_i^2 \sigma_i^2$
	<p>$Z=\max, \min(X_1, X_2, \dots, X_n)$</p>	<p>若 $X_1, X_2 \dots X_n$ 相互独立, 其分布函数分别为 $F_{x_1}(x), F_{x_2}(x) \dots F_{x_n}(x)$, 则 $Z=\max, \min(X_1, X_2, \dots, X_n)$ 的分布函数为:</p> $F_{\max}(x) = F_{x_1}(x) \bullet F_{x_2}(x) \dots F_{x_n}(x)$ $F_{\min}(x) = 1 - [1 - F_{x_1}(x)] \bullet [1 - F_{x_2}(x)] \dots [1 - F_{x_n}(x)]$

	χ^2 分布	<p>设 n 个随机变量 X_1, X_2, \dots, X_n 相互独立, 且服从标准正态分布, 可以证明它们的平方和</p> $W = \sum_{i=1}^n X_i^2$ <p>服从自由度为 n 的 χ^2 分布, 记为 $W \sim \chi^2(n)$。</p> <p>所谓自由度是指独立正态随机变量的个数, 它是随机变量分布中的一个重要参数。</p> <p>χ^2 分布满足可加性: 设</p> $Y_i \sim \chi^2(n_i),$ <p>则</p> $Z = \sum_{i=1}^k Y_i \sim \chi^2(n_1 + n_2 + \dots + n_k).$
	t 分布	<p>设 X, Y 是两个相互独立的随机变量, 且</p> $X \sim N(0,1), Y \sim \chi^2(n),$ <p>可以证明函数</p> $T = \frac{X}{\sqrt{Y/n}}$ <p>服从自由度为 n 的 t 分布, 记为 $T \sim t(n)$。</p> $t_{1-\alpha}(n) = -t_{\alpha}(n)$
	F 分布	<p>设 $X \sim \chi^2(n_1), Y \sim \chi^2(n_2)$, 且 X 与 Y 独立, 可以证明</p> $F = \frac{X/n_1}{Y/n_2}$ <p>服从第一个自由度为 n_1, 第二个自由度为 n_2 的 F 分布, 记为 $F \sim f(n_1, n_2)$。</p> $F_{1-\alpha}(n_1, n_2) = \frac{1}{F_{\alpha}(n_2, n_1)}$

第四章 随机变量的数字特征

(1)		离散型	连续型
-----	--	-----	-----

一 维 随 机 变 量 的 数 字 特 征	期望 期望就是平均值	设 X 是离散型随机变量, 其分布律为 $P(X = x_k) = p_k$, $k=1,2,\cdots,n$, $E(X) = \sum_{k=1}^n x_k p_k$ (要求绝对收敛)	设 X 是连续型随机变量, 其概率密度为 $f(x)$, $E(X) = \int_{-\infty}^{+\infty} x f(x) dx$ (要求绝对收敛)
	函数的期望	$Y=g(X)$ $E(Y) = \sum_{k=1}^n g(x_k) p_k$	$Y=g(X)$ $E(Y) = \int_{-\infty}^{+\infty} g(x) f(x) dx$
	方差 $D(X)=E[X-E(X)]^2$, 标准差 $\sigma(X) = \sqrt{D(X)}$,	$D(X) = \sum_k [x_k - E(X)]^2 p_k$	$D(X) = \int_{-\infty}^{+\infty} [x - E(X)]^2 f(x) dx$
	矩	①对于正整数 k , 称随机变量 X 的 k 次幂的数学期望为 X 的 k 阶原点矩, 记为 v_k , 即 $v_k = E(X^k) = \sum_i x_i^k p_i,$ $k=1,2, \cdots$. ②对于正整数 k , 称随机变量 X 与 $E(X)$ 差的 k 次幂的数学期望为 X 的 k 阶中心矩, 记为 μ_k , 即 $\mu_k = E(X - E(X))^k$ $= \sum_i (x_i - E(X))^k p_i,$ $k=1,2, \cdots$.	①对于正整数 k , 称随机变量 X 的 k 次幂的数学期望为 X 的 k 阶原点矩, 记为 v_k , 即 $v_k = E(X^k) = \int_{-\infty}^{+\infty} x^k f(x) dx,$ $k=1,2, \cdots$. ②对于正整数 k , 称随机变量 X 与 $E(X)$ 差的 k 次幂的数学期望为 X 的 k 阶中心矩, 记为 μ_k , 即 $\mu_k = E(X - E(X))^k$ $= \int_{-\infty}^{+\infty} (x - E(X))^k f(x) dx,$ $k=1,2, \cdots$.
	切比雪夫不等式	设随机变量 X 具有数学期望 $E(X) = \mu$, 方差 $D(X) = \sigma^2$, 则对于任意正数 ε , 有下列切比雪夫不等式 $P(X - \mu \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$ 切比雪夫不等式给出了在未知 X 的分布的情况下, 对概率 $P(X - \mu \geq \varepsilon)$ 的一种估计, 它在理论上具有重要意义。	

(2) 期 望 的 性 质	(1) $E(C)=C$ (2) $E(CX)=CE(X)$ (3) $E(X+Y)=E(X)+E(Y)$, $E(\sum_{i=1}^n C_i X_i) = \sum_{i=1}^n C_i E(X_i)$ (4) $E(XY)=E(X)E(Y)$, 充分条件: X 和 Y 独立; 充要条件: X 和 Y 不相关。		
(3) 方 差 的 性 质	(1) $D(C)=0$; $E(C)=C$ (2) $D(aX)=a^2D(X)$; $E(aX)=aE(X)$ (3) $D(aX+b)=a^2D(X)$; $E(aX+b)=aE(X)+b$ (4) $D(X)=E(X^2)-E^2(X)$ (5) $D(X \pm Y)=D(X)+D(Y)$, 充分条件: X 和 Y 独立; 充要条件: X 和 Y 不相关。 $D(X \pm Y)=D(X)+D(Y) \pm 2E[(X-E(X))(Y-E(Y))]$, 无条件成立。 而 $E(X+Y)=E(X)+E(Y)$, 无条件成立。		
(4) 常 见 分 布 的 期 望 和 方 差		期望	方差
	0-1 分布 $B(1, p)$	p	$p(1-p)$
	二项分布 $B(n, p)$	np	$np(1-p)$
	泊松分布 $P(\lambda)$	λ	λ
	几何分布 $G(p)$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
	超几何分布 $H(n, M, N)$	$\frac{nM}{N}$	$\frac{nM}{N} \left(1 - \frac{M}{N}\right) \left(\frac{N-n}{N-1}\right)$
	均匀分布 $U(a, b)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
	指数分布 $e(\lambda)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
	正态分布 $N(\mu, \sigma^2)$	μ	σ^2
	χ^2 分布	n	2n
	t 分布	0	$\frac{n}{n-2}$ (n>2)

(5) 二 维 随 机 变 量 的 数 字 特 征	期望	$E(X) = \sum_{i=1}^n x_i p_{i\bullet}$ $E(Y) = \sum_{j=1}^n y_j p_{\bullet j}$	$E(X) = \int_{-\infty}^{+\infty} x f_X(x) dx$ $E(Y) = \int_{-\infty}^{+\infty} y f_Y(y) dy$
	函数的期望	$E[G(X, Y)] =$ $\sum_i \sum_j G(x_i, y_j) p_{ij}$	$E[G(X, Y)] =$ $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} G(x, y) f(x, y) dx dy$
	方差	$D(X) = \sum_i [x_i - E(X)]^2 p_{i\bullet}$ $D(Y) = \sum_j [y_j - E(Y)]^2 p_{\bullet j}$	$D(X) = \int_{-\infty}^{+\infty} [x - E(X)]^2 f_X(x) dx$ $D(Y) = \int_{-\infty}^{+\infty} [y - E(Y)]^2 f_Y(y) dy$
	协方差	<p>对于随机变量 X 与 Y，称它们的二阶混合中心矩 μ_{11} 为 X 与 Y 的协方差或相关矩，记为 σ_{XY} 或 $\text{cov}(X, Y)$，即</p> $\sigma_{XY} = \mu_{11} = E[(X - E(X))(Y - E(Y))].$ <p>与记号 σ_{XY} 相对应，X 与 Y 的方差 D (X) 与 D (Y) 也可分别记为 σ_{XX} 与 σ_{YY}。</p>	

	相关系数	<p>对于随机变量 X 与 Y, 如果 $D(X) > 0, D(Y) > 0$, 则称</p> $\frac{\sigma_{XY}}{\sqrt{D(X)}\sqrt{D(Y)}}$ <p>为 X 与 Y 的相关系数, 记作 ρ_{XY} (有时可简记为 ρ)。</p> <p>$\rho \leq 1$, 当 $\rho = 1$ 时, 称 X 与 Y 完全相关: $P(X = aY + b) = 1$</p> <p>完全相关 $\begin{cases} \text{正相关, 当 } \rho = 1 \text{ 时 } (a > 0), \\ \text{负相关, 当 } \rho = -1 \text{ 时 } (a < 0), \end{cases}$</p> <p>而当 $\rho = 0$ 时, 称 X 与 Y 不相关。</p> <p>以下五个命题是等价的:</p> <p>① $\rho_{XY} = 0$;</p> <p>② $\text{cov}(X, Y) = 0$;</p> <p>③ $E(XY) = E(X)E(Y)$;</p> <p>④ $D(X+Y) = D(X) + D(Y)$;</p> <p>⑤ $D(X-Y) = D(X) + D(Y)$。</p>
	协方差矩阵	$\begin{pmatrix} \sigma_{XX} & \sigma_{XY} \\ \sigma_{YX} & \sigma_{YY} \end{pmatrix}$
	混合矩	<p>对于随机变量 X 与 Y, 如果有 $E(X^k Y^l)$ 存在, 则称之为 X 与 Y 的 $k+l$ 阶混合原点矩, 记为 ν_{kl}; $k+l$ 阶混合中心矩记为:</p> $u_{kl} = E[(X - E(X))^k (Y - E(Y))^l].$
(6) 协方差的性质	<p>(i) $\text{cov}(X, Y) = \text{cov}(Y, X)$;</p> <p>(ii) $\text{cov}(aX, bY) = ab \text{cov}(X, Y)$;</p> <p>(iii) $\text{cov}(X_1 + X_2, Y) = \text{cov}(X_1, Y) + \text{cov}(X_2, Y)$;</p> <p>(iv) $\text{cov}(X, Y) = E(XY) - E(X)E(Y)$。</p>	
(7) 独立和不相关	<p>(i) 若随机变量 X 与 Y 相互独立, 则 $\rho_{XY} = 0$; 反之不真。</p> <p>(ii) 若 $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, 则 X 与 Y 相互独立的充要条件是 X 和 Y 不相关。</p>	

第五章 大数定律和中心极限定理

<p>(1) 大数定律</p> $\bar{X} \rightarrow \mu$	<p>切比雪夫大数定律</p>	<p>设随机变量 X_1, X_2, \dots 相互独立, 均具有有限方差, 且被同一常数 C 所界: $D(X_i) < C (i=1, 2, \dots)$, 则对于任意的正数 ε, 有</p> $\lim_{n \rightarrow \infty} P\left(\left \frac{1}{n} \sum_{i=1}^n X_i - \frac{1}{n} \sum_{i=1}^n E(X_i)\right < \varepsilon\right) = 1.$ <p>特殊情形: 若 X_1, X_2, \dots 具有相同的数学期望 $E(X_i) = \mu$, 则上式成为</p> $\lim_{n \rightarrow \infty} P\left(\left \frac{1}{n} \sum_{i=1}^n X_i - \mu\right < \varepsilon\right) = 1.$
	<p>伯努利大数定律</p>	<p>设 μ 是 n 次独立试验中事件 A 发生的次数, p 是事件 A 在每次试验中发生的概率, 则对于任意的正数 ε, 有</p> $\lim_{n \rightarrow \infty} P\left(\left \frac{\mu}{n} - p\right < \varepsilon\right) = 1.$ <p>伯努利大数定律说明, 当试验次数 n 很大时, 事件 A 发生的频率与概率有较大判别的可能性很小, 即</p> $\lim_{n \rightarrow \infty} P\left(\left \frac{\mu}{n} - p\right \geq \varepsilon\right) = 0.$ <p>这就以严格的数学形式描述了频率的稳定性。</p>
	<p>辛钦大数定律</p>	<p>设 $X_1, X_2, \dots, X_n, \dots$ 是相互独立同分布的随机变量序列, 且 $E(X_n) = \mu$, 则对于任意的正数 ε 有</p> $\lim_{n \rightarrow \infty} P\left(\left \frac{1}{n} \sum_{i=1}^n X_i - \mu\right < \varepsilon\right) = 1.$
<p>(2) 中心极限定理</p> $\bar{X} \rightarrow N\left(\mu, \frac{\sigma^2}{n}\right)$	<p>列维-林德伯格定理</p>	<p>设随机变量 X_1, X_2, \dots 相互独立, 服从同一分布, 且具有相同的数学期望和方差:</p> $E(X_k) = \mu, D(X_k) = \sigma^2 \neq 0 (k=1, 2, \dots),$ <p>则随机变量</p> $Y_n = \frac{\sum_{k=1}^n X_k - n\mu}{\sqrt{n}\sigma}$ <p>的分布函数 $F_n(x)$ 对任意的实数 x, 有</p> $\lim_{n \rightarrow \infty} F_n(x) = \lim_{n \rightarrow \infty} P\left\{\frac{\sum_{k=1}^n X_k - n\mu}{\sqrt{n}\sigma} \leq x\right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt.$ <p>此定理也称为独立同分布的中心极限定理。</p>

	棣莫弗-拉普拉斯定理	<p>设随机变量 X_n 为具有参数 $n, p(0 < p < 1)$ 的二项分布, 则对于任意实数 x, 有</p> $= \lim_{n \rightarrow \infty} P \left\{ \frac{X_n - np}{\sqrt{np(1-p)}} \leq x \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt.$
(3) 泊松定理		<p>若当 $n \rightarrow \infty$ 时, $np \rightarrow \lambda > 0$, 则</p> $C_n^k p^k (1-p)^{n-k} \rightarrow \frac{\lambda^k}{k!} e^{-\lambda} \quad (n \rightarrow \infty).$ <p>其中 $k=0, 1, 2, \dots, n, \dots$。 二项分布的极限分布为泊松分布。</p>

第六章 样本及抽样分布

(1) 数理统计的基本概念	总体	在数理统计中, 常把被考察对象的某一个 (或多个) 指标的全体称为总体 (或母体)。我们总是把总体看成一个具有分布的随机变量 (或随机向量)。
	个体	总体中的每一个单元称为样品 (或个体)。
	样本	<p>我们把从总体中抽取的部分样品 x_1, x_2, \dots, x_n 称为样本。样本中所含的样品数称为样本容量, 一般用 n 表示。在一般情况下, 总是把样本看成是 n 个相互独立的且与总体有相同分布的随机变量, 这样的样本称为简单随机样本。在泛指任一次抽取的结果时, x_1, x_2, \dots, x_n 表示 n 个随机变量 (样本); 在具体的一次抽取之后, x_1, x_2, \dots, x_n 表示 n 个具体的数值 (样本值)。我们称之为样本的两重性。</p>
	样本函数和统计量	<p>设 x_1, x_2, \dots, x_n 为总体的一个样本, 称</p> $\varphi = \varphi(x_1, x_2, \dots, x_n)$ <p>为样本函数, 其中 φ 为一个连续函数。如果 φ 中不包含任何未知参数, 则称 $\varphi(x_1, x_2, \dots, x_n)$ 为一个统计量。</p>

	<p>常见统计量及其性质</p>	<p>样本均值 $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i.$</p> <p>样本方差 $S^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2.$</p> <p>样本标准差 $S = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}.$</p> <p>样本 k 阶原点矩 $M_k = \frac{1}{n} \sum_{i=1}^n x_i^k, k=1,2,\dots.$</p> <p>样本 k 阶中心矩 $M'_k = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^k, k=2,3,\dots.$</p> <p>$E(\bar{X}) = \mu, D(\bar{X}) = \frac{\sigma^2}{n},$</p> <p>$E(S^2) = \sigma^2, E(S^{*2}) = \frac{n-1}{n} \sigma^2,$</p> <p>其中 $S^{*2} = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$, 为二阶中心矩。</p>
<p>(2) 正态总体下的四大分布</p>	<p>正态分布</p>	<p>设 x_1, x_2, \dots, x_n 为来自正态总体 $N(\mu, \sigma^2)$ 的一个样本, 则样本函数</p> $u \stackrel{\text{def}}{=} \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} \sim N(0,1).$
	<p>t 分布</p>	<p>设 x_1, x_2, \dots, x_n 为来自正态总体 $N(\mu, \sigma^2)$ 的一个样本, 则样本函数</p> $t \stackrel{\text{def}}{=} \frac{\bar{x} - \mu}{s / \sqrt{n}} \sim t(n-1),$ <p>其中 $t(n-1)$ 表示自由度为 $n-1$ 的 t 分布。</p>

	χ^2 分布	<p>设 x_1, x_2, \dots, x_n 为来自正态总体 $N(\mu, \sigma^2)$ 的一个样本, 则样本函数</p> $w \stackrel{\text{def}}{=} \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1),$ <p>其中 $\chi^2(n-1)$ 表示自由度为 $n-1$ 的 χ^2 分布。</p>
	F分布	<p>设 x_1, x_2, \dots, x_{n_1} 为来自正态总体 $N(\mu, \sigma_1^2)$ 的一个样本, 而 y_1, y_2, \dots, y_{n_2} 为来自正态总体 $N(\mu, \sigma_2^2)$ 的一个样本, 则样本函数</p> $F \stackrel{\text{def}}{=} \frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2} \sim F(n_1-1, n_2-1),$ <p>其中</p> $S_1^2 = \frac{1}{n_1-1} \sum_{i=1}^{n_1} (x_i - \bar{x})^2, \quad S_2^2 = \frac{1}{n_2-1} \sum_{i=1}^{n_2} (y_i - \bar{y})^2;$ <p>$F(n_1-1, n_2-1)$ 表示第一自由度为 n_1-1, 第二自由度为 n_2-1 的 F 分布。</p>
(3) 正态总体下分布的性质	\bar{X} 与 S^2 独立。	

	极大似然估计	<p>当总体 X 为连续型随机变量时，设其分布密度为 $f(x; \theta_1, \theta_2, \dots, \theta_m)$，其中 $\theta_1, \theta_2, \dots, \theta_m$ 为未知参数。又设 x_1, x_2, \dots, x_n 为总体的一个样本，称</p> $L(\theta_1, \theta_2, \dots, \theta_m) = \prod_{i=1}^n f(x_i; \theta_1, \theta_2, \dots, \theta_m)$ <p>为样本的似然函数，简记为 L_n。</p> <p>当总体 X 为离散型随机变量时，设其分布律为 $P\{X = x\} = p(x; \theta_1, \theta_2, \dots, \theta_m)$，则称</p> $L(x_1, x_2, \dots, x_n; \theta_1, \theta_2, \dots, \theta_m) = \prod_{i=1}^n p(x_i; \theta_1, \theta_2, \dots, \theta_m)$ <p>为样本的似然函数。</p> <p>若似然函数 $L(x_1, x_2, \dots, x_n; \theta_1, \theta_2, \dots, \theta_m)$ 在 $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_m$ 处取到最大值，则称 $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_m$ 分别为 $\theta_1, \theta_2, \dots, \theta_m$ 的最大似然估计值，相应的统计量称为最大似然估计量。</p> $\left. \frac{\partial \ln L_n}{\partial \theta_i} \right _{\theta_i = \hat{\theta}_i} = 0, i = 1, 2, \dots, m$ <p>若 $\hat{\theta}$ 为 θ 的极大似然估计，$g(x)$ 为单调函数，则 $g(\hat{\theta})$ 为 $g(\theta)$ 的极大似然估计。</p>
(2) 估计量的评选标准	无偏性	<p>设 $\hat{\theta} = \hat{\theta}(x_1, x_2, \dots, x_n)$ 为未知参数 θ 的估计量。若 $E(\hat{\theta}) = \theta$，则称 $\hat{\theta}$ 为 θ 的无偏估计量。</p> $E(\bar{X}) = E(X), \quad E(S^2) = D(X)$
	有效性	<p>设 $\hat{\theta}_1 = \hat{\theta}_1(x_1, x_2, \dots, x_n)$ 和 $\hat{\theta}_2 = \hat{\theta}_2(x_1, x_2, \dots, x_n)$ 是未知参数 θ 的两个无偏估计量。若 $D(\hat{\theta}_1) < D(\hat{\theta}_2)$，则称 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 有效。</p>

	一致性	<p>设 $\hat{\theta}_n$ 是 θ 的一串估计量, 如果对于任意的正数 ε, 都有</p> $\lim_{n \rightarrow \infty} P(\hat{\theta}_n - \theta > \varepsilon) = 0,$ <p>则称 $\hat{\theta}_n$ 为 θ 的一致估计量 (或相合估计量)。</p> <p>若 $\hat{\theta}$ 为 θ 的无偏估计, 且 $D(\hat{\theta}) \rightarrow 0 (n \rightarrow \infty)$, 则 $\hat{\theta}$ 为 θ 的一致估计。</p> <p>只要总体的 $E(X)$ 和 $D(X)$ 存在, 一切样本矩和样本矩的连续函数都是相应总体的一致估计量。</p>
(3) 区间估计	置信区间和置信度	<p>设总体 X 含有一个待估的未知参数 θ。如果我们从样本 x_1, x_2, \dots, x_n 出发, 找出两个统计量 $\theta_1 = \theta_1(x_1, x_2, \dots, x_n)$ 与 $\theta_2 = \theta_2(x_1, x_2, \dots, x_n)$ ($\theta_1 < \theta_2$), 使得区间 $[\theta_1, \theta_2]$ 以 $1 - \alpha$ ($0 < \alpha < 1$) 的概率包含这个待估参数 θ, 即</p> $P\{\theta_1 \leq \theta \leq \theta_2\} = 1 - \alpha,$ <p>那么称区间 $[\theta_1, \theta_2]$ 为 θ 的置信区间, $1 - \alpha$ 为该区间的置信度 (或置信水平)。</p>
	单正态总体的期望和方差的区间估计	<p>设 x_1, x_2, \dots, x_n 为总体 $X \sim N(\mu, \sigma^2)$ 的一个样本, 在置信度为 $1 - \alpha$ 下, 我们来确定 μ 和 σ^2 的置信区间 $[\theta_1, \theta_2]$。具体步骤如下:</p> <p>(i) 选择样本函数;</p> <p>(ii) 由置信度 $1 - \alpha$, 查表找分位数;</p> <p>(iii) 导出置信区间 $[\theta_1, \theta_2]$。</p>
	已知方差, 估计均值	<p>(i) 选择样本函数</p> $u = \frac{\bar{x} - \mu}{\sigma_0 / \sqrt{n}} \sim N(0, 1).$ <p>(ii) 查表找分位数</p> $P\left(-\lambda \leq \frac{\bar{x} - \mu}{\sigma_0 / \sqrt{n}} \leq \lambda\right) = 1 - \alpha.$ <p>(iii) 导出置信区间</p> $\left[\bar{x} - \lambda \frac{\sigma_0}{\sqrt{n}}, \bar{x} + \lambda \frac{\sigma_0}{\sqrt{n}}\right]$

	未知方差，估计均值	<p>(i) 选择样本函数</p> $t = \frac{\bar{x} - \mu}{S / \sqrt{n}} \sim t(n-1).$ <p>(ii) 查表找分位数</p> $P\left(-\lambda \leq \frac{\bar{x} - \mu}{S / \sqrt{n}} \leq \lambda\right) = 1 - \alpha.$ <p>(iii) 导出置信区间</p> $\left[\bar{x} - \lambda \frac{S}{\sqrt{n}}, \bar{x} + \lambda \frac{S}{\sqrt{n}}\right]$
	方差的区间估计	<p>(i) 选择样本函数</p> $w = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1).$ <p>(ii) 查表找分位数</p> $P\left(\lambda_1 \leq \frac{(n-1)S^2}{\sigma^2} \leq \lambda_2\right) = 1 - \alpha.$ <p>(iii) 导出 σ 的置信区间</p> $\left[\sqrt{\frac{n-1}{\lambda_2}} S, \sqrt{\frac{n-1}{\lambda_1}} S\right]$

第八章 假设检验

基本思想	<p>假设检验的统计思想是，概率很小的事件在一次试验中可以认为基本上是不会发生的，即小概率原理。</p> <p>为了检验一个假设 H_0 是否成立。我们先假定 H_0 是成立的。如果根据这个假定导致了一个不合理的事件发生，那就表明原来的假定 H_0 是不正确的，我们拒绝接受 H_0；如果由此没有导出不合理的现象，则不能拒绝接受 H_0，我们称 H_0 是相容的。与 H_0 相对的假设称为备择假设，用 H_1 表示。</p> <p>这里所说的小概率事件就是事件 $\{K \in R_\alpha\}$，其概率就是检验水平 α，通常我们取 $\alpha=0.05$，有时也取 0.01 或 0.10。</p>
基本步骤	<p>假设检验的基本步骤如下：</p> <p>(i) 提出零假设 H_0；</p> <p>(ii) 选择统计量 K；</p> <p>(iii) 对于检验水平 α 查表找分位数 λ；</p> <p>(iv) 由样本值 x_1, x_2, \dots, x_n 计算统计量之值 K；</p> <p>将 \hat{K} 与 λ 进行比较，作出判断：当 $\hat{K} > \lambda$ (或 $\hat{K} > \lambda$) 时否定 H_0，否则认为 H_0 相容。</p>

两类错误	第一类错误	当 H_0 为真时，而样本值却落入了否定域，按照我们规定的检验法则，应当否定 H_0 。这时，我们把客观上 H_0 成立判为 H_0 为不成立（即否定了真实的假设），称这种错误为“以真当假”的错误或第一类错误，记 α 为犯此类错误的概率，即 $P\{\text{否定 } H_0 H_0 \text{ 为真}\} = \alpha$ ； 此处的 α 恰好为检验水平。
	第二类错误	当 H_1 为真时，而样本值却落入了相容域，按照我们规定的检验法则，应当接受 H_0 。这时，我们把客观上 H_0 不成立判为 H_0 成立（即接受了不真实的假设），称这种错误为“以假当真”的错误或第二类错误，记 β 为犯此类错误的概率，即 $P\{\text{接受 } H_0 H_1 \text{ 为真}\} = \beta$ 。
	两类错误的关系	人们当然希望犯两类错误的概率同时都很小。但是，当容量 n 一定时， α 变小，则 β 变大；相反地， β 变小，则 α 变大。取定 α 要想使 β 变小，则必须增加样本容量。 在实际使用时，通常人们只能控制犯第一类错误的概率，即给定显著性水平 α 。 α 大小的选取应根据实际情况而定。当我们宁可“以假为真”、而不愿“以真当假”时，则应把 α 取得很小，如 0.01，甚至 0.001。反之，则应把 α 取得大些。

单正态总体均值和方差的假设检验

条件	零假设	统计量	对应样本函数分布	否定域
已知 σ^2	$H_0: \mu = \mu_0$	$U = \frac{\bar{x} - \mu_0}{\sigma_0 / \sqrt{n}}$	$N(0, 1)$	$ u > u_{1-\frac{\alpha}{2}}$
	$H_0: \mu \leq \mu_0$			$u > u_{1-\alpha}$
	$H_0: \mu \geq \mu_0$			$u < -u_{1-\alpha}$
未知 σ^2	$H_0: \mu = \mu_0$	$T = \frac{\bar{x} - \mu_0}{S / \sqrt{n}}$	$t(n-1)$	$ t > t_{1-\frac{\alpha}{2}}(n-1)$
	$H_0: \mu \leq \mu_0$			$t > t_{1-\alpha}(n-1)$
	$H_0: \mu \geq \mu_0$			$t < -t_{1-\alpha}(n-1)$
未知 σ^2	$H_0: \sigma^2 = \sigma_0^2$	$w = \frac{(n-1)S^2}{\sigma_0^2}$	$\chi^2(n-1)$	$w < \chi_{\frac{\alpha}{2}}^2(n-1)$ 或 $w > \chi_{1-\frac{\alpha}{2}}^2(n-1)$

	$H_0: \sigma^2 \leq \sigma_0^2$			$w > \kappa_{1-a}^2(n-1)$
	$H_0: \sigma^2 \geq \sigma_0^2$			$w < \kappa_a^2(n-1)$