概率论与数理统计

孙旭磊 2023年2月

第一章 事件及概率

事件的运算

当事件不相容时，可以把"并"运算符号改写为通常的加号

基本模型

1. 从 n 个不同的元素中，有放回地取 r 个元素组成的组合（不考虑顺序），不同方式个数为 C_{n+r-1}^r （重复组合数）
2. 把 r 个球随机放到不同编号的 n 个盒子中去，球不可辨，每个盒中不限球的个数，不同的排法个数为 C_{n+r-1}^r
3. 把 n 个不同的元素分为有序的 k 个部分，第 i 部分有 r_i 个元素，$i=1,2,\cdots,k$，$r_1+r_2+\cdots+r_k=n$，则不同的分法个数为 $\dfrac{n!}{r_1!r_2!\cdots r_k!}$ （多项式系数）
4. 有 n 个元素，属于 k 个不同的类，同类元素之间不可辨认，各类元素分别有 $n_1,n_2,\cdots,n_k$ 个，其中 $n_1+n_2+\cdots+n_k=n$，要把它们排成一列，则一共有 $\dfrac{n!}{n_1!n_2!\cdots n_k!}$ 种不同的排法

概率的公理化定义

1. 非负性：$0\le P(A)\le 1$
2. 规范性：设 Ω 为必然事件，则 $P(\Omega)=1$
3. 可数可加性：对 Ω 中两两不相容事件 $A_1,A_2,\cdots,A_k\cdots$，有 $P\left(\bigcup\limits_{k=1}^{\infty}A_k\right)=\sum\limits_{k=1}^{\infty}P(A_k)$

概率的性质

1. $P(\varnothing)=0$
2. 有限可加性：若 $A_k,k=1,2,\cdots,n$ 且两两不相容，则 $P\left(\sum\limits_{k=1}^{n}A_k\right)=\sum\limits_{k=1}^{n}P(A_k)$
3. 可减性：若 $A\subset B$，则 $P(B-A)=P(B)-P(A)$
4. 单调性：若 $A\subset B$，则 $P(A)\le P(B)$
5. $P(\bar{A})=1-P(A)$
6. 容斥原理：
$$P\left(\bigcup_{k=1}^{n}A_k\right)=\sum_{k=1}^{n}P(A_k)-\sum_{1\le i\le j\le n}P(A_iA_j)+\sum_{1\le i\le j\le k\le n}P(A_iA_jA_k)-\cdots+(-1)^{n-1}P(A_1A_2\cdots A_n)$$
7. 次可加性：对任意的事件列 $A_1,A_2,\cdots,A_n,\cdots$，有 $P\left(\bigcup\limits_{k=1}^{\infty}A_k\right)\le\sum\limits_{k=1}^{\infty}P(A_k)$

条件概率

- $P(A|B)=\dfrac{P(AB)}{P(B)}$
- 对 Ω 中两两不相容事件 $A_1,A_2,\cdots,A_k\cdots$，有 $P\left(\bigcup\limits_{k=1}^{\infty}A_k|B\right)=\sum\limits_{k=1}^{\infty}P(A_k|B)$
- 设 $B_1,B_2,\cdots,B_n$ 是 Ω 的一个完备事件群，A 为 Ω 中任一事件，则有
  - 全概率公式：$P(A)=\sum P(A|B_i)P(B_i)$
  - 贝叶斯公式：$P(B_i|A)=\dfrac{P(B_iA)}{P(A)}=\dfrac{P(A|B_i)P(B_i)}{\sum P(A|B_j)P(B_j)}$

事件的独立性

- A,B 相互独立 $\Leftrightarrow P(AB)=P(A)P(B)\Leftrightarrow P(B|A)=P(B)$（若 $P(A)>0$）
- 两两独立 $\Leftarrow$ 相互独立，两两独立 $\nRightarrow$ 相互独立

第二章 随机变量及其分布

分布	分布律/概率密度函数	期望	方差	性质
0-1 分布	$\begin{pmatrix} 0 & 1 \\ 1-p & p \end{pmatrix}$	p	$p(1-p)$	
二项分布 $B(n,p)$	$C_n^k p^k(1-p)^{n-k}$	np	$np(1-p)$	再生性
超几何分布	$\dfrac{C_M^m C_{N-M}^{n-m}}{C_N^n}$	$\dfrac{nM}{N}$	$\dfrac{nM(N-M)(N-n)}{N^2(N-1)}$	
几何分布 $G(p)$	$p(1-p)^{k-1}$	$\dfrac{1}{p}$	$\dfrac{1-p}{p^2}$	无记忆性
负二项分布 $NB(r,p)$	$C_{k-1}^{r-1}p^r(1-p)^{k-r}$	$\dfrac{r}{p}$	$\dfrac{r(1-p)}{p^2}$	
泊松分布 $P(\lambda)$	$\dfrac{\lambda^k}{k!}e^{-\lambda}$	λ	λ	再生性
连续均匀分布 $U(a,b)$	$\dfrac{1}{b-a}I_{(a,b)}(x)$	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$	
指数分布 $\text{Exp}(\lambda)$	$\lambda e^{-\lambda x}I_{(0,\infty)}(x)$	$\dfrac{1}{\lambda}$	$\dfrac{1}{\lambda^2}$	无记忆性
正态分布 $N(\mu,\sigma^2)$	$\dfrac{1}{\sqrt{2\pi}\sigma}\exp\left[-\dfrac{(x-\mu)^2}{2\sigma^2}\right]$	μ	σ	再生性

离散型分布

1. 0-1 分布（伯努利分布/两点分布）
2. 离散均匀分布
3. 二项分布 $B(n,p)$
   - 再生性：独立随机变量 $X\sim B(n,p),Y\sim B(m,p)\Rightarrow X+Y\sim B(n+m,p)$

4. 超几何分布：池塘中有 N 条鱼，捞 M 条鱼标记后再放回，现从中捞出 n 条鱼，恰有 m 条鱼有标记的概率（不放回抽样）
5. 几何分布 $G(p)$：伯努利实验成功前（含）的次数
   - 无记忆性：$P(X>m+n|X>m)=P(X>n)=(1-p)^n$
6. 负二项分布（帕斯卡分布）$NB(r,p)$：伯努利实验第 k 次时恰成功了 r 次
7. 泊松分布 $P(\lambda)$
   - 泊松逼近定理：设 $X_n\sim B(n,p_n)$，若当 $n\to\infty$ 时，$np_n\to\lambda>0$，则 $\lim\limits_{n\to\infty}P(X_n=k)=\dfrac{\lambda^k}{k!}e^{-\lambda}$
   - 再生性：独立随机变量 $X\sim P(\lambda),Y\sim P(\mu)\Rightarrow X+Y\sim P(\lambda+\mu)$

连续型分布

1. 连续均匀分布 $U(a,b)$
2. 指数分布 $\mathrm{Exp}(\lambda)$
   - 无记忆性：$P(X>s+t|X>t)=P(X>s)$
3. 正态分布 $N(\mu,\sigma^2)$
   - 标准正态分布：$\varphi(x)=\dfrac{1}{\sqrt{2\pi}}\exp\left(-\dfrac{x^2}{2}\right)$
   - $\mathrm{E}\left(X^2\right)=\mu^2+\sigma^2,\mathrm{E}(|X-\mu|)=\sqrt{\dfrac{2}{\pi}}\sigma$
   - 标准化变换：$\dfrac{X-\mu}{\sigma}\sim N(0,1)$，$F(x)=\Phi\left(\dfrac{x-\mu}{\sigma}\right)$
   - 再生性：独立随机变量 $X\sim N(\mu_1,\sigma_1^2),Y\sim N(\mu_2,\sigma_2^2)\Rightarrow X+Y\sim N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$
   - 二元正态分布：$(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，
     $f(x,y)=\dfrac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left\{-\dfrac{1}{2(1-\rho^2)}\left[\dfrac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\dfrac{x-\mu_1}{\sigma_1}\dfrac{y-\mu_2}{\sigma_2}+\dfrac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\}$,
     $-1<\rho<1$；$C_1X+C_2Y\sim N(C_1\mu_1+C_2\mu_2,C_1^2\sigma_1^2+C_2^2\sigma_2^2+2C_1C_2\sigma_1\sigma_2\rho)$

分布函数与概率密度函数

- $F(x)=P(X\le x)$，右连续（$F(x+0)=F(x)$）
- $F(x)=\displaystyle\int_{-\infty}^{x}f(t)\mathrm{d}t$，$X\sim f(x)$

连续型随机变量函数的分布

- 设 $X\sim f(x),Y=g(X)$，则随机变量 Y 的分布函数为
  $F_1(y)=P(Y\le y)=P(g(X)\le y)=\displaystyle\int_{g(x)\le y}f(x)\mathrm{d}x$
- 若 $g(x)$ 是严格单调的且反函数 $h(y)$ 可导，则有概率密度函数 $f_1(y)=\begin{cases}f(h(y))|h'(y)|, & \alpha<y<\beta\\ 0, & \text{其他}\end{cases}$，其中 $\alpha=\min\{g(-\infty),g(+\infty)\},\beta=\max\{g(-\infty),g(+\infty)\}$

第三章 多维随机变量及其分布

联合分布函数与联合密度函数

若 $f(x,y)$ 在 (x_0,y_0) 处连续，则 $\left.\dfrac{\partial^2 F(x,y)}{\partial x\partial y}\right|_{(x_0,y_0)} = f(x_0,y_0)$

边缘分布

- $F_1(x)=F(x,\infty),F_2(y)=F(\infty,y)$
- $f_1(x)=\int_{-\infty}^{+\infty} f(x,y)\mathrm{d}y, f_2(y)=\int_{-\infty}^{+\infty} f(x,y)\mathrm{d}x$

条件分布

- 离散型：$p_{i|j}=\dfrac{p_{ij}}{p_{\cdot j}}$
- 连续型：$f_{X|Y}(x|y)=\dfrac{f(x,y)}{f_2(y)}, f_2(y)>0$

随机变量的独立性

X,Y 相互独立 $\Leftrightarrow F(x,y)=F_1(x)F_2(y)\Leftrightarrow$

- 离散型：$p_{ij}=p_{i\cdot}p_{\cdot j}$
- 连续型：$f(x,y)=f_1(x)f_2(y)\Leftrightarrow f(x,y)=g_1(x)g_2(y)$ 即 $f(x,y)$ 可分离变量

随机向量函数的分布

- 设 $(X,Y)\sim f(x,y), Z=g(X,Y)$，则 $F_Z(z)=P(Z\le z)=\iint_{g(x,y)\le z} f(x,y)\mathrm{d}x\mathrm{d}y$
- 设 $Z_1=g_1(X,Y),Z_2=g_2(X,Y),u=g_1(x,y),v=g_2(x,y),x=\varphi_1(u,v),y=\varphi_2(u,v),\mathbf{J}=\begin{vmatrix}\frac{\partial x}{\partial u} & \frac{\partial x}{\partial v}\\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v}\end{vmatrix}$，则 $F_Z(z_1,z_2)=\iint_{g_1(x,y)\le z_1,g_2(x,y)\le z_2} f(x,y)\mathrm{d}x\mathrm{d}y=\iint_{u\le z_1,v\le z_2} f(\varphi_1(u,v),\varphi_2(u,v))|\mathbf{J}|\mathrm{d}u\mathrm{d}v$，$f_Z(z_1,z_2)=f(\varphi_1(z_1,z_2),\varphi_2(z_1,z_2))|\mathbf{J}||_{(u,v)=(z_1,z_2)}$
- 随机变量和的分布：$f_{X+Y}(z)=\int_{-\infty}^{+\infty} f(x,z-x)\mathrm{d}x \xlongequal{\text{独立}} \int_{-\infty}^{+\infty} f_1(x)f_2(z-x)\mathrm{d}x\equiv f_1*f_2(z)$
- 随机变量积的分布：$f_{XY}(z)=\int_{-\infty}^{+\infty}\left|\frac{1}{t}\right| f\left(t,\frac{z}{t}\right)\mathrm{d}t$
- 随机变量商的分布：$f_{\frac{X}{Y}}(z)=\int_{-\infty}^{+\infty}|t|f(zt,t)\mathrm{d}t$
- 独立随机变量最大（小）值的分布：$F_{\max}(z)=F_1(z)F_2(z),F_{\min}(z)=1-(1-F_1(z))(1-F_2(z))$

第四章 随机变量的数学特征和极限定理

期望

期望存在的条件

- 离散型：$\sum_{i=1}^{\infty}|x_i|p_i<\infty$
- 连续型：$\int_{-\infty}^{+\infty}|x|f(x)\mathrm{d}x<\infty$
- 先判断存在性，再计算期望

期望的性质

1. 线性性：（无条件）$\mathrm{E}(\sum X_i)=\sum \mathrm{E}(X_i)$
2. 对于两两不相关随机变量，$\mathrm{E}(\prod X_i)=\prod \mathrm{E}(X_i)$
3. $\mathrm{E}(g(X))=\int_{-\infty}^{+\infty}g(x)f(x)\mathrm{d}x$

条件期望

- 离散型：$\mathrm{E}(Y|X=x_k)=\sum y_iP(Y=y_i|X=x_k)$
- 连续型：$\mathrm{E}(Y|X=x)=\int_{-\infty}^{+\infty}yf(y|x)\mathrm{d}y$
- $\mathrm{E}[\mathrm{E}(Y|X)]=\mathrm{E}(Y)$

中位数

- $P(X\le m)\ge\frac{1}{2},P(X\ge m)\ge\frac{1}{2}$，不一定唯一
- $\min_a \mathrm{E}\left(|x-a|\right)=\mathrm{E}\left(|x-m|\right)$
- p 分位数 Q_p：$P(X\le Q_p)\ge p$，$P(X\ge Q_p)\ge 1-p$

方差和标准差

定义

方差：$\mathrm{Var}(X)=\mathrm{E}\left((X-\mathrm{E}(X))^2\right)$

标准差：$\sigma=\sqrt{\mathrm{Var}(X)}$

方差的性质

1. $\mathrm{Var}(X)=\mathrm{E}\left(X^2\right)-(\mathrm{E}(X))^2$（常用此式计算）
2. $\mathrm{Var}(c_1X+c_2)=c_1^2\,\mathrm{Var}(X)$
3. 对于两两不相关随机变量，$\mathrm{Var}\left(\sum_{i=1}^{n}X_i\right)=\sum_{i=1}^{n}\mathrm{Var}\left(X_i\right)$，$\mathrm{Var}\left(\sum_{i=1}^{n}c_iX_i\right)=\sum_{i=1}^{n}c_i^2\,\mathrm{Var}\left(X_i\right)$；对于独立同分布，$\mathrm{Var}\left(\sum_{i=1}^{n}\frac{1}{n}X_i\right)=\mathrm{Var}(\overline{X})=\frac{\sigma^2}{n}$
4. $\mathrm{Var}(X)\le\mathrm{E}\left((X-c)^2\right)$，等号成立当且仅当 $c=\mathrm{E}(X)$

5. 方差分解公式：$\operatorname{Var}(X)=\operatorname{Var}[\mathrm{E}(X|Y)]+\mathrm{E}[\operatorname{Var}(X|Y)]$

随机变量标准化

$Y=\dfrac{X-\mathrm{E}(X)}{\sqrt{\operatorname{Var}(X)}}$，则 $\mathrm{E}(Y)=0,\operatorname{Var}(Y)=1$

矩

- X 关于 c 点的 k 阶矩：$\mathrm{E}\left((X-c)^k\right)$
  - k 阶原点矩：$a_k=\mathrm{E}(X^k)$
  - k 阶中心矩：$\mu_k=\mathrm{E}\left((X-\mathrm{E}(X))^k\right)$
- 矩母函数：$M_X(s)=\mathrm{E}\left(e^{Xs}\right)$
  - 存在的条件：$\exists a>0$，使得 $\forall s\in[-a,a]$，$M_X(s)$ 是有限的
  - $M_X(s)=\displaystyle\sum_{k=0}^{\infty}\mathrm{E}\left(X^k\right)\frac{s^k}{k!}$
  - $\mathrm{E}(X^k)=\dfrac{\mathrm{d}^k}{\mathrm{d}s^k}M_X(s)\Big|_{s=0}$
  - 若 $\exists a>0$，使得 $\forall s\in[-a,a]$，$M_X(s),M_Y(s)$ 有限且相等，则 $\forall t\in\mathbb{R},F_X(t)=F_Y(t)$

协方差和相关系数

定义

- 协方差：$\operatorname{Cov}(X,Y)=\mathrm{E}[(X-\mathrm{E}(X))(Y-\mathrm{E}(Y))]$
- 相关系数：$\rho_{X,Y}=\dfrac{\operatorname{Cov}(X,Y)}{\sqrt{\operatorname{Var}(X)\operatorname{Var}(Y)}}$

协方差的性质

1. $\operatorname{Cov}(X,Y)=\operatorname{Cov}(Y,X)$
2. $\operatorname{Cov}(X,X)=\operatorname{Var}(X)$
3. $\operatorname{Cov}(X,Y)=\mathrm{E}(XY)-\mathrm{E}(X)\,\mathrm{E}(Y)$（常用此式计算）
4. 双线性：$\operatorname{Cov}(aX+bY,cX+dY)=ac\operatorname{Var}(X)+(ad+bc)\operatorname{Cov}(X,Y)+bd\operatorname{Var}(Y)$
5. 若 X 与 Y 互相独立，则 $\operatorname{Cov}(X,Y)=0$
6. 柯西-施瓦茨不等式：$(\operatorname{Cov}(X,Y))^2\le\operatorname{Var}(X)\operatorname{Var}(Y)$，等号成立当且仅当 $c_1X+c_2Y+c_3=0$

随机变量的不相关

- X,Y 相互独立 $\Rightarrow\rho=0$，反之不一定成立；若 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，则 X,Y 相互独立 $\Leftrightarrow\rho=0$
- X,Y 不相关 $\Leftrightarrow\operatorname{Cov}(X,Y)=0\Leftrightarrow\mathrm{E}(XY)=\mathrm{E}(X)\,\mathrm{E}(Y)\Leftrightarrow\operatorname{Var}(X+Y)=\operatorname{Var}(X)+\operatorname{Var}(Y)$

熵

- 离散型：$H(X)=-\sum_{k=1}^{\infty}p_k\log_2(p_k)$
- 连续型：$H(X)=-\int_{-\infty}^{\infty}f_X(x)\ln f_X(x)\mathrm{d}x$

马尔可夫不等式和切比雪夫不等式

- 马尔可夫不等式：若随机变量 $Y\geq 0$，则 $\forall\varepsilon>0$，有 $P(Y\geq\varepsilon)\leq\dfrac{\mathrm{E}(Y)}{\varepsilon}$
- 切比雪夫不等式：随机变量 X，$\forall\varepsilon>0$，有 $P(|X-\mu|\geq\varepsilon)\leq\dfrac{\mathrm{Var}(X)}{\varepsilon^2}$

依概率和依分布收敛

- 依概率收敛 ($X_n\xrightarrow{P}X$)：$\forall\varepsilon>0,\lim\limits_{n\to\infty}P(|X_n-X|\geq\varepsilon)=0$
- 依分布收敛/弱收敛 ($X_n\xrightarrow{\mathcal{L}}X$)：对 F 的所有连续点 x, $\lim\limits_{n\to\infty}F_n(x)=F(x)$
- $X_n\xrightarrow{P}X\Rightarrow X_n\xrightarrow{\mathcal{L}}X$，$X_n\xrightarrow{\mathcal{L}}c\Rightarrow X_n\xrightarrow{P}c$

大数定律和中心极限定理

- 大数定律：设随机变量 $X_1,X_2,\cdots$ 独立同分布，则 $\forall\varepsilon>0,\lim\limits_{n\to\infty}P\left(\left|\overline{X}_n-\mu\right|\geqslant\varepsilon\right)=0$
- 伯努利大数定律：设 $\{X_k\}$ 为独立的 0-1 分布随机变量序列，$P(X_k=1)=p$，则 $\dfrac{1}{n}\sum\limits_{k=1}^{n}X_k\xrightarrow{P}p$
- 林德伯格-莱维中心极限定理：设随机变量 $X_1,X_2,\cdots$ 独立同分布，则 $\lim\limits_{n\to\infty}P\left(\dfrac{\overline{X}_n-\mu}{\sigma/\sqrt{n}}\leqslant x\right)=\Phi(x)$
- 棣莫弗-拉普拉斯中心极限定理：设 $\{X_k\}$ 为独立的 0-1 分布随机变量序列，$P(X_k=1)=p$，$S_n=\sum\limits_{k=1}^{n}X_k$，

  则 $\lim\limits_{n\to\infty}P\left(\dfrac{S_n-np}{\sqrt{np(1-p)}}\leqslant x\right)=\Phi(x)$

第五章 统计学基本概念

基本概念

- 简单随机抽样：同分布且独立
- 统计量：由样本 $X=(X_1,X_2,\cdots,X_n)$ 决定的量，且只与样本有本，与任何未知参数无关
- 样本方差：$S^2=\dfrac{1}{n-1}\sum\limits_{i=1}^{n}(X_i-\overline{X})^2$
- 次序统计量 $X_{(1)},\cdots,X_{(n)}$：把样本排序后得到 $X_{(1)}\leq X_{(2)}\leq\cdots\leq X_{(n)}$；极差：$X_{(n)}-X_{(1)}$

- **抽样分布**

1. χ^2 分布 χ_n^2：设 $X_1, X_2, \cdots, X_n$ 是来自标准正态总体 $N(0,1)$ 的一个样本，则统计量 $X = X_1^2 + X_2^2 + \cdots + X_n^2$ 服从自由度为 n 的 χ^2 分布
   - $k_n(x) = \dfrac{1}{\Gamma(\frac{n}{2})2^{\frac{n}{2}}} e^{-\frac{x}{2}} x^{\frac{n-2}{2}} I_{(0,\infty)}(x)$
   - $\mathrm{E}(X) = n, \operatorname{Var}(X) = 2n$
   - 再生性：独立随机变量 $X \sim \chi_m^2, Y \sim \chi_n^2 \Rightarrow X + Y \sim \chi_{m+n}^2$
2. t 分布 t_n：设独立随机变量 $X \sim N(0,1), Y \sim \chi_n^2$，则 $T = \dfrac{X}{\sqrt{Y/n}}$ 服从自由度为 n 的 t 分布
   - $n \geq 2$ 时 $\mathrm{E}(T) = 0$；$n \geq 3$ 时 $\operatorname{Var}(T) = \dfrac{n}{n-2}$
   - $n = 1$ 时化为柯西分布，$f_1(t) = \dfrac{1}{\pi(1+t^2)}$
   - $n \to \infty$ 时，$t_n \to N(0,1)$
3. F 分布 $F_{m,n}$：设独立随机变量 $X \sim \chi_m^2, Y \sim \chi_n^2$，则 $F = \dfrac{X/m}{Y/n}$ 服从自由度为 m, n 的 F 分布
   - $F \sim F_{m,n} \Rightarrow 1/F \sim F_{n,m}$
   - $T \sim t_n \Rightarrow T^2 \sim F_{1,n}$
   - $F_{m,n}(1-\alpha) = 1/F_{n,m}(\alpha)$（$F(\alpha)$：上 α 分位数）

- 设随机变量 $X_1, X_2, \cdots, X_n$ i.i.d. $\sim N(\mu, \sigma^2)$，则 $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2$，$\overline{X}$ 和 S^2 相互独立，$\dfrac{\overline{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}$
- 设 $X_1, X_2, \cdots, X_m$ 是来自正态总体 $N\left(\mu_1, \sigma_1^2\right)$ 的样本，样本 $Y_1, Y_2, \cdots, Y_n$ 来自正态总体 $N\left(\mu_2, \sigma_2^2\right)$，且 X 与 Y 相互独立，则
  - $\dfrac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F_{m-1,n-1}$
  - 若 $\sigma_1^2 = \sigma_2^2$，则 $\dfrac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{S_T\sqrt{\dfrac{1}{m}+\dfrac{1}{n}}} \sim t_{m+n-2}$，其中 $S_T = \sqrt{\dfrac{(m-1)S_1^2+(n-1)S_2^2}{m+n-2}}$
- Γ 函数
  - $\Gamma(s) = \displaystyle\int_0^{+\infty} t^{s-1} e^{-t} \mathrm{d}t$
  - $\Gamma(s+1) = s\Gamma(s)$，$\Gamma(n+1) = n!$
  - $\Gamma(\frac{1}{2}) = \sqrt{\pi}$，$\Gamma(n+\frac{1}{2}) = \dfrac{(2n-1)!!}{2^n}\sqrt{\pi}$
  - $\Gamma(s)\Gamma(1-s) = \dfrac{\pi}{\sin s\pi}$
  - $\Gamma(2s) = \dfrac{2^{2s-1}}{\sqrt{\pi}}\Gamma(s)\Gamma\left(s+\dfrac{1}{2}\right)$

第六章 参数点估计

矩估计 $\hat{\theta}_M$

样本矩：$a_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k, m_k=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^k$

总体矩：$\alpha_k(\theta)=\mathrm{E}\left(X^k\right),\mu_k(\theta)=\mathrm{E}\left((X-\mathrm{E}(X))^k\right)$

$a_k\xrightarrow{P}\alpha_k, m_k\xrightarrow{P}\mu_k$，故用 a_k,m_k 估计 α_k,μ_k（尽量使用低阶矩），解方程得 $\hat{\theta}_i(X_1,\cdots,X_n)$，作为 θ 的估计

函数的矩估计：用 $g(\hat{\theta})$ 估计 $g(\theta)$

最大似然估计(MLE) $\hat{\theta}_L$

似然函数 $L(\theta)$ 或 $L(\theta;x)$：等于联合概率密度函数 $f(x;\theta)$

最大似然估计：对于固定的 x，需满足 $L(\hat{\theta})=\max\limits_{\theta\in\Theta}L(\theta)$

对数似然函数 $\ell(\theta)=\ln L(\theta)$：$\frac{\partial\ell(\theta)}{\partial\theta_i}=\frac{\partial L(\theta)}{\partial\theta_i}=0$

优良性准则

- 无偏估计量：设 $\hat{g}(X_1,\cdots,X_n)$ 为待估计参数函数 $g(\theta)$ 的一个估计量，$\forall\theta\in\Theta,\mathrm{E}_\theta[\hat{g}(X_1,\cdots,X_n)]=g(\theta)$
- 均方误差 $\mathrm{MSE}_\theta(\hat{\theta})=\mathrm{E}_\theta[(\hat{\theta}-\theta)^2]=\text{波动}+\text{偏差}=\mathrm{Var}_\theta(\hat{\theta})+[\mathrm{E}_\theta(\hat{\theta})-\theta]^2$
- 有效性：设 $\hat{\theta}_1,\hat{\theta}_2$ 都是总体参数 θ 的无偏估计，若 $\mathrm{Var}_\theta(\hat{\theta}_1)\leqslant\mathrm{Var}_\theta(\hat{\theta}_2),\forall\theta\in\Theta$，且至少存在一个 θ 使不等号成立，则 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 更有效
- 相合估计量：$\hat{\theta}(X_1,\cdots,X_n)\xrightarrow{P}\theta$
- 渐近正态性：记 $\mathrm{Var}_\theta(\hat{\theta}(X_1,\cdots,X_n))=\sigma_n^2(\theta)$，有 $\frac{\hat{\theta}(X_1,\cdots,X_n)-\theta}{\sigma_n(\theta)}\xrightarrow{\mathcal{L}}N(0,1)$，即 $\lim\limits_{n\to\infty}P\left(\frac{\hat{\theta}(X_1,\cdots,X_n)-\theta}{\sigma_n(\theta)}\le x\right)=\Phi(x),\forall x$

第七章 区间估计

基本概念

- 设 $\hat{\theta}_1,\hat{\theta}_2$ 为样本 $X_1,\cdots,X_n$ 的统计量，$P_\theta(\hat{\theta}_1\le\theta\le\hat{\theta}_2)=1-\alpha,\forall\theta\in\Theta$，$[\hat{\theta}_1,\hat{\theta}_2]$ 为置信区间估计，置信系数为 $1-\alpha$
- 上 α 分位数 w_α：$F(w_\alpha)=1-\alpha$；下 α 分位数 v_α：$F(v_\alpha)=F(w_{1-\alpha})=\alpha$
- 置信上限：$P_\theta(\overline{\theta}\ge\theta)=1-\alpha,\forall\theta\in\Theta$；置信下限：$P_\theta(\underline{\theta}\le\theta)=1-\alpha,\forall\theta\in\Theta$

枢轴变量法

1. 找到一个 θ 的良好点估计 $T(X)$，一般为最大似然估计
2. 构造枢轴变量 $S(T,\theta)$，使得它的分布 F 已知，不能包含其他未知参数
3. 要求 $\forall a<b$，$a\le S(T,\theta)\le b$ 可改写为 $A\le\theta\le B$，其中 A,B 只能与 $T(X),a,b$ 有关，与 θ 无关
4. $P(w_{1-\alpha/2}\le S(T,\theta)\le w_{\alpha/2})=1-\alpha\Rightarrow A\le\theta\le B$

正态总体参数的区间估计

均值 μ

σ 已知：枢轴变量 $\dfrac{\sqrt{n}(\overline{X}-\mu)}{\sigma} \sim N(0,1)$，置信区间 $\overline{x} \pm \dfrac{\sigma}{\sqrt{n}} u_{\alpha/2}$

σ 未知：枢轴变量 $\dfrac{\sqrt{n}(\overline{X}-\mu)}{S} \sim t_{n-1}$，置信区间 $\overline{x} \pm \dfrac{s}{\sqrt{n}} t_{n-1}(\alpha/2)$

方差 σ^2

μ 已知：枢轴变量 $\dfrac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2 \sim \chi_n^2$，置信区间 $\left[\dfrac{\sum_{i=1}^n (X_i-\mu)^2}{\chi_n^2(\alpha/2)}, \dfrac{\sum_{i=1}^n (X_i-\mu)^2}{\chi_n^2(1-\alpha/2)}\right]$

μ 未知：枢轴变量 $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2$，置信区间 $\left[\dfrac{(n-1)s^2}{\chi_{n-1}^2(\alpha/2)}, \dfrac{(n-1)s^2}{\chi_{n-1}^2(1-\alpha/2)}\right]$

两个正态总体均值差 $\mu_2 - \mu_1$

σ_1^2, σ_2^2 已知：枢轴变量 $\dfrac{(\overline{Y}-\overline{X})-(\mu_2-\mu_1)}{\sqrt{\dfrac{\sigma_1^2}{m}+\dfrac{\sigma_2^2}{n}}} \sim N(0,1)$，置信区间 $(\overline{y}-\overline{x}) \pm \sqrt{\dfrac{\sigma_1^2}{m}+\dfrac{\sigma_2^2}{n}}\, u_{\alpha/2}$

σ_1^2, σ_2^2 未知：若 $\sigma_1^2=\sigma_2^2$，记 $S_T=\sqrt{\dfrac{(m-1)S_1^2+(n-1)S_2^2}{m+n-2}}$，枢轴变量 $\dfrac{(\overline{Y}-\overline{X})-(\mu_2-\mu_1)}{S_T\sqrt{\dfrac{1}{m}+\dfrac{1}{n}}} \sim t_{m+n-2}$，

置信区间 $\overline{y}-\overline{x} \pm \sqrt{\dfrac{1}{m}+\dfrac{1}{n}}\, s_T t_{m+n-2}(\alpha/2)$

两个正态总体方差比 σ_1^2/σ_2^2

枢轴变量 $\dfrac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F_{m-1,n-1}$，置信区间 $\left[\dfrac{s_1^2}{s_2^2}F_{n-1,m-1}(1-\alpha/2), \dfrac{s_1^2}{s_2^2}F_{n-1,m-1}(\alpha/2)\right]$

0-1 分布总体比例 p 的区间估计

每次发生概率 p，n 次实验，发生 Y_n 次；$n \gg 1$

枢轴变量 $\dfrac{Y_n-np}{\sqrt{np(1-p)}} \sim N(0,1)$，置信区间 $\dfrac{\hat{p}+\dfrac{u_{\alpha/2}^2}{2n}}{1+\dfrac{u_{\alpha/2}^2}{n}} \pm u_{\alpha/2}\dfrac{\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}+\dfrac{u_{\alpha/2}^2}{4n^2}}}{1+\dfrac{u_{\alpha/2}^2}{n}}$（得分区间），其中点估计

$\hat{p}=\dfrac{y_n}{n}$，瓦尔德置信区间 $\hat{p} \pm u_{\alpha/2}\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$（要求 $n\hat{p}>10, n(1-\hat{p})>10$，可忽略 $\dfrac{u_{\alpha/2}^2}{n}$）

一般总体均值 μ 的区间估计

若 $n \gg 1$，枢轴变量 $\frac{\sqrt{n}(\overline{X}-\mu)}{S} \sim N(0,1)$，置信区间 $\overline{x} \pm \frac{s}{\sqrt{n}} u_{\alpha/2}$

第八章 假设检验

基本概念

- 检验函数：$\Psi(X)=\begin{cases}1, & X \in \text{拒绝域 } D \\ 0, & X \in \text{接受域 } A\end{cases}$
- 功效函数：$\beta_\Psi(\theta)=P_\theta(\text{在检验 } \Psi \text{ 下 } H_0 \text{ 被否定})$
  - 当真实参数 θ^* 属于 H_0 时，希望 $\beta_\Psi(\theta^*)$ 尽量小；反之亦然
  - 若 $\forall \theta \in H_0, \beta_\Psi(\theta) \leq \alpha$，则称 Ψ 为 H_0 的一个水平 α 的检验，即为错误拒绝 H_0 所允许的最大概率
- 两类错误
  - 第一类（弃真错误）：$\alpha_{1\Psi}(\theta)=\begin{cases}\beta_\Psi(\theta), & \theta \in H_0 \\ 0, & \theta \in H_1\end{cases}$
  - 第二类（存伪错误）：$\alpha_{2\Psi}(\theta)=\begin{cases}0, & \theta \in H_0 \\ 1-\beta_\Psi(\theta), & \theta \in H_1\end{cases}$

显著性检验的步骤

1. 求出未知参数 θ 的一个较优的点估计 $\hat{\theta}$，如最大似然估计
2. 寻找检验统计量 $T=T(\hat{\theta},\theta_0)$ 使得当 $\theta=\theta_0$ 时，T 的分布已知（如 $N(0,1), t_n, F_{m,n}$）
3. 寻找拒绝域，由 $P(H_0\text{被拒绝}|\theta \in \Theta_0) \leq \alpha$ 得到临界值
4. 判断检验统计量是否落在拒绝域中

正态总体参数的检验

均值 μ

σ 已知：检验统计量 $U=\frac{\sqrt{n}(\overline{X}-\mu_0)}{\sigma} \sim N(0,1)$

- $H_0: \mu=\mu_0 \leftrightarrow H_1: \mu \neq \mu_0$：拒绝域 $|U|>u_{\alpha/2}$
- $H_0: \mu \leq \mu_0 \leftrightarrow H_1: \mu>\mu_0$：拒绝域 $U>u_\alpha$
- $H_0: \mu \geq \mu_0 \leftrightarrow H_1: \mu<\mu_0$：拒绝域 $U<-u_\alpha$

σ 未知：检验统计量 $T=\frac{\sqrt{n}(\overline{X}-\mu_0)}{S} \sim t_{n-1}$

- $H_0: \mu=\mu_0 \leftrightarrow H_1: \mu \neq \mu_0$：拒绝域 $|T|>t_{n-1}(\alpha/2)$
- $H_0: \mu \leq \mu_0 \leftrightarrow H_1: \mu>\mu_0$：拒绝域 $T>t_{n-1}(\alpha)$
- $H_0: \mu \geq \mu_0 \leftrightarrow H_1: \mu<\mu_0$：拒绝域 $T<-t_{n-1}(\alpha)$

方差 σ^2

μ 已知：检验统计量 $\chi^2 = \frac{1}{\sigma_0^2}\sum_{i=1}^{n}(X_i-\mu)^2 \sim \chi_n^2$

- $H_0:\sigma^2=\sigma_0^2 \leftrightarrow H_1:\sigma^2\neq\sigma_0^2$：拒绝域 $\chi^2>\chi_n^2(\alpha/2)$ & $\chi^2<\chi_n^2(1-\alpha/2)$
- $H_0:\sigma^2\leq\sigma_0^2 \leftrightarrow H_1:\sigma^2>\sigma_0^2$：拒绝域 $\chi^2>\chi_n^2(\alpha)$
- $H_0:\sigma^2\geq\sigma_0^2 \leftrightarrow H_1:\sigma^2<\sigma_0^2$：拒绝域 $\chi^2<\chi_n^2(1-\alpha)$

μ 未知：检验统计量 $\chi^2 = \frac{1}{\sigma_0^2}\sum_{i=1}^{n}(X_i-\overline{X})^2 \sim \chi_{n-1}^2$

- $H_0:\sigma^2=\sigma_0^2 \leftrightarrow H_1:\sigma^2\neq\sigma_0^2$：拒绝域 $\chi^2>\chi_{n-1}^2(\alpha/2)$ & $\chi^2<\chi_{n-1}^2(1-\alpha/2)$
- $H_0:\sigma^2\leq\sigma_0^2 \leftrightarrow H_1:\sigma^2>\sigma_0^2$：拒绝域 $\chi^2>\chi_{n-1}^2(\alpha)$
- $H_0:\sigma^2\geq\sigma_0^2 \leftrightarrow H_1:\sigma^2<\sigma_0^2$：拒绝域 $\chi^2<\chi_{n-1}^2(1-\alpha)$

两个正态总体均值差 $\mu_1-\mu_2$

成组比较

$(X_1,\cdots,X_m)\sim N(\mu_1,\sigma),(Y_1,\cdots,Y_n)\sim N(\mu_2,\sigma)$

σ 已知：检验统计量 $U=\dfrac{\overline{X}-\overline{Y}}{\sigma\sqrt{\frac{1}{m}+\frac{1}{n}}}\sim N(0,1)$

- $H_0:\mu_1=\mu_2 \leftrightarrow H_1:\mu_1\neq\mu_2$：拒绝域 $|U|>u_{\alpha/2}$
- $H_0:\mu_1\leq\mu_2 \leftrightarrow H_1:\mu_1>\mu_2$：拒绝域 $U>u_\alpha$
- $H_0:\mu_1\geq\mu_2 \leftrightarrow H_1:\mu_1<\mu_2$：拒绝域 $U<-u_\alpha$

σ 未知：检验统计量 $T=\dfrac{\overline{X}-\overline{Y}}{S_T\sqrt{\frac{1}{m}+\frac{1}{n}}}\sim t_{m+n-2}$

- $H_0:\mu_1=\mu_2 \leftrightarrow H_1:\mu_1\neq\mu_2$：拒绝域 $|T|>t_{m+n-2}(\alpha/2)$
- $H_0:\mu_1\leq\mu_2 \leftrightarrow H_1:\mu_1>\mu_2$：拒绝域 $T>t_{m+n-2}(\alpha)$
- $H_0:\mu_1\geq\mu_2 \leftrightarrow H_1:\mu_1<\mu_2$：拒绝域 $T<-t_{m+n-2}(\alpha/2)$

成对比较

$Z=Y-X,Z_i=Y_i-X_i$，化归为单个正态总体的情形

两个正态总体方差比 σ_1^2/σ_2^2

μ 未知：检验统计量 $F=\dfrac{S_1^2}{S_2^2}\sim F_{m-1,n-1}$

- $H_0:\sigma_1^2=\sigma_2^2 \leftrightarrow H_1:\sigma_1^2\neq\sigma_2^2$：拒绝域 $F>F_{m-1,n-1}(\alpha/2)$ & $F<F_{m-1,n-1}(1-\alpha/2)$
- $H_0:\sigma_1^2\leq\sigma_2^2 \leftrightarrow H_1:\sigma_1^2>\sigma_2^2$：拒绝域 $F>F_{m-1,n-1}(\alpha)$
- $H_0:\sigma_1^2\geq\sigma_2^2 \leftrightarrow H_1:\sigma_1^2<\sigma_2^2$：拒绝域 $F<F_{m-1,n-1}(1-\alpha)$

- **0-1 分布总体比例 p 的检验**

$$X=\sum_{i=1}^{n}X_i\sim B(n,p)$$

$H_0:p\le p_0\leftrightarrow H_1:p>p_0$：功效函数 $\beta_\Psi(p_0)=1-\sum_{i=0}^{C}\binom{n}{i}p_0^i(1-p_0)^{n-i}=1-F_{p_0}(C)=\alpha$

- **p 值**

p 值 $=P($得到当前样本下检验统计量的值或更极端的值$|$原假设下$)$

当 p 值 $<\alpha$ 时，拒绝原假设 H_0

第九章 非参数假设检验

- **理论分布完全已知且只取有限个值**

$H_0:P(X=a_i)=p_i,i=1,2,\cdots,k\leftrightarrow H_1:\exists j,\text{s.t.}P(X=a_j)\ne p_j$

理论值 $E_i=np_i$，观测值 $O_i=n_i$

检验统计量 $Z=\sum\frac{(O-E)^2}{E}=\sum_{i=1}^{k}\frac{(np_i-n_i)^2}{np_i}=\sum_{i=1}^{k}\frac{n_i^2}{np_i}-n\overset{n\to\infty}{\sim}\chi^2_{k-1}$，拒绝域 $Z>\chi^2_{k-1}(\alpha)$

拟合优度 $p(Z_0)=P(Z\ge Z_0)=1-F_{\chi^2_{k-1}}(Z_0)$

- **理论分布类型已知但含有有限个未知参数**

$H_0:P(X=a_i)=p_i(\theta_1,\cdots,\theta_r),i=1,\cdots,k$

检验统计量 $Z=\sum\frac{(O-\hat{E})^2}{\hat{E}}=\sum_{i=1}^{n}\frac{n_i^2}{n\hat{p}_i}-n\overset{n\to\infty}{\sim}\chi^2_{k-r-1}$，拒绝域 $Z>\chi^2_{k-r-1}(\alpha)$

- **列联表检验**

H_0：属性 A,B 独立

检验统计量 $Z=\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{(n_{ij}-n_{i\cdot}n_{\cdot j}/n)^2}{n_{i\cdot}n_{\cdot j}/n}=\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{(nn_{ij}-n_{i\cdot}n_{\cdot j})^2}{nn_{i\cdot}n_{\cdot j}}\overset{n\to\infty}{\sim}\chi^2_{(a-1)(b-1)}$

拒绝域 $Z>\chi^2_{(a-1)(b-1)}(\alpha)$