知识小料

「电计 2203 班」周常规知识整理共享

25

日期：2024-4-22
学科：概率与统计 A

本文档用于对《概率与统计 A》课程作出简明复习。

1 常见的分布

二项分布 $X \sim B(n,p)$

$$P(X=k) = C_n^k p^k (1-p)^{n-k}$$

$E(X) = np$

$D(X) = np(1-p)$

特殊性质：无

泊松分布 $X \sim P(\lambda)$

$$P(X=k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

$E(X) = \lambda$

$D(X) = \lambda$

特殊性质：泊松逼近定理

几何分布 $X \sim G(p)$

$$P(X=k) = (1-p)^{k-1} \cdot p$$

$E(X) = \frac{1}{p}$

$D(X) = \frac{1-p}{p^2}$

特殊性质：无记忆性

均匀分布 $X \sim U(a,b)$

$$f(x) = \frac{1}{b-a} (a<x<b)$$

$E(X) = \frac{a+b}{2}$

$D(X) = \frac{(b-a)^2}{12}$

特殊性质：无

指数分布 $X \sim e(\theta)$

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

$E(X) = \frac{1}{\theta}$

$D(X) = \frac{1}{\theta^2}$

特殊性质：无记忆性；极小分布

$W = \min\{X_i\} \sim e(n\theta)$

正态分布 $X \sim N(\mu,\sigma^2)$

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$E(X) = \mu$

$D(X) = \sigma^2$

特殊性质：太多了

χ^2 分布——$\chi^2 \sim \chi^2(n)$

$(X_1, X_2, \ldots, X_n) \sim N(0,1)$ 独立

$$\chi^2 = X_1^2 + X_2^2 + \cdots + X_n^2$$

特殊性质：$E(\chi^2) = n$，$D(\chi^2) = 2n$，$\chi^2(n) + \chi^2(m) \sim \chi^2(n+m)$

t 分布——$t \sim t(n)$

$X \sim N(0,1)$，$Y \sim \chi^2(n)$ 独立

$$t = \frac{X}{\sqrt{\frac{Y}{n}}}$$

特殊性质：$t^2 \sim F(1,n)$

F 分布——$F \sim F(n,m)$

$X \sim \chi^2(n)$，$Y \sim \chi^2(m)$ 独立

$$F = \frac{\frac{X}{n}}{\frac{Y}{m}}$$

特殊性质：$\frac{1}{F} \sim F(m,n)$

「实例 12」

$X_1, X_2, \ldots, X_n$ 相互独立同分布，$E(X_i) = \mu$，$D(X_i) = \sigma^2$，

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

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

$E(\bar{X}) = \mu$，$D(\bar{X}) = \frac{\sigma^2}{n}$

$E(S^2) = \sigma^2$

2 概率论基本概念

2.1 随机事件与运算

样本空间 Ω，样本点 ω，随机事件 $A, B \ldots$，随机变量 $X, Y \ldots$
随机事件关系：

- 子事件：$A \subset B$
- 并事件：$C = A + B = A \cup B$
- 交事件：$C = AB = A \cap B$
- 差事件：$C = A - B = A\bar{B}$
- 互不相容：$AB = \varnothing$
- 分配律、德摩根律

概率定义的三个基本方面：非负性、归一性、可列可加性。

- 减法公式：$P(A-B) = P(A\bar{B}) = P(A) - P(AB)$
- 加法公式：$P(A+B) = P(A) + P(B) - P(AB)$

- 单调性：$B \subset A \implies P(B) \leqslant P(A)$
- $A = \varnothing \implies P(A) = 0$，$B = \Omega \implies P(B) = 1$，但反之不然

古典概型：$P(A) = \dfrac{n_A}{n_\Omega} = \dfrac{A\text{的样本点数}}{\text{样本点总数}}$

几何概型：$P(A) = \dfrac{|A|}{|\Omega|} = \dfrac{A\text{的度量}}{\text{全部度量}}$

相关实例：摸球问题、抽签问题、生日问题等，先定位后挑选；分组法应用

2.2 条件概率与公式

条件概率：$P(A|B) = \dfrac{P(AB)}{P(B)}$

乘法公式：$P(AB) = P(B) \cdot P(A|B)$

全概率公式：设 $A_1, A_2, \ldots$ 是 Ω 的剖分，$P(A_i) > 0$，对任意 B：

$$P(B) = \sum_i P(A_i) \cdot P(B|A_i) \tag{1}$$

贝叶斯公式：

$$P(A_k|B) = \frac{P(A_k) \cdot P(B|A_k)}{\sum_i P(A_i) \cdot P(B|A_i)} \tag{2}$$

2.3 事件独立性

事件独立性的原初定义：$P(AB) = P(A) \cdot P(B)$

互不相容：$P(A+B) = P(A) + P(B)$

多个事件相互独立定义（略）

关于事件独立性…

1. $A_1, A_2, \ldots$ 独立，则对它们做任意种类运算后得到的事件仍然相互独立。

2. X, Y 独立：

   - $P(X = x_i, Y = y_j) = P(X = x_i) \cdot P(Y = y_j)$
   - $P(X \leqslant x, Y \leqslant y) = P(X \leqslant x) \cdot P(Y \leqslant y)$
   - $F(x, y) = F_X(x) \cdot F_Y(y)$
   - $f(x, y) = f_X(x) \cdot f_Y(y)$

3. X, Y 独立：

- $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho) \iff \rho=0$

4. X,Y 独立：
   - $g(x),h(y)$ 为连续函数 $\iff g(X),h(Y)$ 独立

5. X,Y 独立：
   - $E(XY)=E(X)E(Y)$
   - $D(X+Y)=D(X)+D(Y)$
   - $\mathrm{Cov}(X,Y)=0$
   - $\rho_{XY}=0$（注：互不相关不一定相互独立）

3 一维、二维随机变量及其分布

3.1 一维随机变量

随机变量：从样本空间 Ω 到实数集 $\mathbb{R}$ 的映射，记为 $X=X(\omega)$。

分布函数：$F(x):=P(X\leqslant x)$（借用离散数学符号 := 表示定义）

- 单调递增：$x_1<x_2 \implies F(x_1)\leqslant F(x_2)$
- 收敛性：$F(-\infty)=0$，$F(+\infty)=1$
- 右连续性：$\lim\limits_{h\to 0^+} F(x+h)=F(x)$，对任意 x

离散型随机变量分布列：$p_i:=P(X=x_i)$（$i=1,2,\ldots$）

- 非负性：$P(X=x_i)\geqslant 0$
- 归一性：$\sum_{i=1}^{\infty} P(X=x_i)=1$

连续性随机变量密度函数：$f(x)$ 满足 $F(x)=\displaystyle\int_{-\infty}^{t} f(t)\mathrm{d}t$（$\forall x\in\mathbb{R}$）

- 非负性：$f(x)\geqslant 0$，对任意 x
- 归一性：$\displaystyle\int_{-\infty}^{+\infty} f(x)\mathrm{d}x=1$
- 求导：若 $f(x)$ 连续，则 $F(x)$ 可微且 $f(x)=F'(x)$

分布函数 $F(x)$ 须写成左闭右开形式，密度函数 $f(x)$ 不要求。

随机变量函数：随机变量 $X:\Omega\to\mathbb{R}$ 与函数 $g:\mathbb{R}\to\mathbb{R}$ 的复合，记为 $Y=g(X)$。

3.2 一维随机变量的其他内容

常见随机变量的分布，详见第一页。

求随机变量函数的密度函数的方法：写取值范围—列平凡情况—解非平凡情况/转化反解—求导

泊松逼近定理的应用：$X \sim B(n,p)$，n 很大，但 np 不太大，通常 $np \leqslant 5$（有时是 $n(1-p) \leqslant 5$），则有

$$P(X=k) = \mathrm{C}_n^k p^k (1-p)^{n-k} \approx \frac{(np)^k}{k!}\mathrm{e}^{-np} \tag{3}$$

即让 $\lambda = np$。

几何分布的无记忆性：$X \sim G(p)$，对任意自然数 m,n 均有 $P(X>n+m|X>n) = P(X>m)$。

指数分布的无记忆性：$X \sim e(\theta)$，对任意 $s,t>0$ 均有 $P(X>s+t|X>s) = P(X>t)$。

正态分布的标准化：$X \sim N(\mu,\sigma^2)$ 经过 $Y = \dfrac{X-\mu}{\sigma}$ 得到 $Y \sim N(0,1)$。

标准正态分布密度函数：

$$\varphi(x) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}} \qquad X \sim N(0,1) \tag{4}$$

分布函数 $\Phi(x)$ 是偶函数，$\Phi(0) = \frac{1}{2}$。

若 $X \sim N(\mu,\sigma^2)$，$Y = aX+b$，则 $Y \sim N(a\mu+b, a^2\sigma^2)$。正态分布标准化 $Y = \frac{X-\mu}{\sigma}$ 是特例，取 $a = \frac{1}{\sigma}$，$b = -\frac{\mu}{\sigma}$。

3.3 二维随机变量

二维随机变量：从 Ω 到 $\mathbb{R}^2$ 的映射，记为 $(X,Y) = (X(\omega), Y(\omega))$。

联合分布函数：$F(x,y) := P(X \leqslant x, Y \leqslant y)$

- 单调递增：给定 y，$x_1 < x_2 \implies F(x_1,y) \leqslant F(x_2,y)$（给定 x 类似）
- 收敛性：$F(-\infty,y) = F(x,-\infty) = F(-\infty,-\infty) = 0$，$F(+\infty,+\infty) = 1$。
- 右连续性：给定 y，$\lim\limits_{h\to 0^+} F(x+h,y) = F(x,y)$（给定 x 类似）
- 矩形法则：设 $x_1 < x_2, y_1 < y_2$，则 $F(x_2,y_2) - F(x_1,y_2) - F(x_2,y_1) + F(x_1,y_1) \geqslant 0$

边缘分布函数：$F_X(x) := F(x,+\infty)$，$F_Y(y) = F(+\infty,y)$

离散型随机变量联合分布列：$p_{ij} := P(X=x_i, Y=y_j)$

- 非负性：$p_{ij} \geqslant 0$
- 归一性：$\sum\limits_i\sum\limits_j p_{ij} = 1$

种类	表达式	简记
联合 (X,Y)	$P(X=x_i,\quad Y=y_j)$	p_{ij}
边缘 X	$P(X=x_i)=P\left(X=x_i,\quad \bigcup_{j=1}^{m}\{Y=y_j\}\right)$	$p_{i\bullet}$
边缘 Y	$P(Y=y_j)=P\left(\bigcup_{i=1}^{n}\{X=x_i\},\quad Y=y_j\right)$	$p_{\bullet j}$

表 1: 二维离散型随机变量的三种分布列（• 相当于通配符）

边缘分布列见表 1。

条件分布列：给定 x_i 及 $p_{i\bullet}>0$，称一列数：$P(Y=y_j|X=x_i)=\dfrac{P(X=x_i,Y=y_j)}{P(X=x_i)}$ 为 $X=x_i$ 发生条件下的 y 的条件分布列。简记：$p_{j|i}=\dfrac{p_{ij}}{p_{i\bullet}}$

连续性随机变量联合密度函数：$f(x,y)$ 满足 $F(x,y)=\displaystyle\int_{-\infty}^{x}\int_{-\infty}^{y}f(u,v)\mathrm{d}u\mathrm{d}v$ $(\forall x,y\in\mathbb{R})$

- 非负性：$f(x,y)\geqslant 0$
- 归一性：$\displaystyle\iint_{\mathbb{R}^2}f(x,y)\mathrm{d}x\mathrm{d}y=1$
- 求偏导：若 f 连续，则 F 二阶可微且 $f(x,y)=F''_{xy}(x,y)$

边缘密度函数见下框。

关于联合确定边缘…

$$f_X(x)=f(x,+\infty)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}y$$

$$f_Y(y)=f(+\infty,y)=\int_{-\infty}^{+\infty}f(x,y)\mathrm{d}x$$

条件密度函数：

对给定 $Y=y$，$f_Y(y)>0$，则 $f_{X|Y}(x|y)=\dfrac{f(x,y)}{f_Y(y)}$ 是 X 的条件密度函数；

对给定 $X=x$，$f_X(x)>0$，则 $f_{Y|X}(y|x)=\dfrac{f(x,y)}{f_X(x)}$ 是 Y 的条件密度函数。

二维随机变量函数：随机变量 $(X,Y):\Omega\to\mathbb{R}^2$ 与二元函数 $g:\mathbb{R}^2\to\mathbb{R}$ 的复合，记为 $Z=g(X,Y)$。

3.4 二维随机变量的其他内容

二维均匀分布：(X,Y)，D 是 $\mathbb{R}^2$ 的有界区域，$|D|$ 为面积。若

$$f(x,y)=\begin{cases}\frac{1}{|D|}, & (x,y)\in D\\ 0, & \text{其他}\end{cases} \tag{5}$$

则称 (X,Y) 服从 D 上的二维均匀分布。

二维正态分布：(X,Y)，若

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\mathrm{e}^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2}-\frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right]} \tag{6}$$

则称 (X,Y) 服从二维正态分布，记为 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$。

相关系数 $\rho=0$（即 X,Y 独立），且均值 $\mu_1=\mu_2=0$，方差 $\sigma_1^2=\sigma_2^2=1$ 时：

$$f(x,y)=\frac{1}{2\pi}\mathrm{e}^{-\frac{x^2+y^2}{2}} \qquad (X,Y)\sim N(0,0,1,1,0) \tag{7}$$

这与一维标准正态分布 $f(x)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}} \qquad X\sim N(0,1)$ 很相似。

X 与 Y 有相同分布列：指 X,Y 取值相同，且取到对应值的概率也相同。

极小分布与指数分布：若 $X_i\sim e(\theta)$ 且相互独立，$W=\min\limits_{1\leqslant i\leqslant n}\{X_i\}$，则 $W\sim e(n\theta)$。

泊松分布可加性：X,Y 独立，$X\sim P(\lambda_1)$，$Y\sim P(\lambda_2)$，则 $Z=X+Y\sim P(\lambda_1+\lambda_2)$。

关于正态分布…

1. 一维线性组合：$X\sim N(\mu_1,\sigma_1^2)$，$Y\sim N(\mu_2,\sigma_2^2)$，且 X,Y 相互独立，则

$$aX+bY+c\sim N(a\mu_1+b\mu_2+c,\quad a^2\sigma_1^2+b^2\sigma_2^2)$$

$X_i\sim N(\mu_i,\sigma_i^2)$（$i=1,2,\ldots,n$），且各个 X_i 相互独立，则

$$b+\sum_{i=1}^n a_iX_i\sim N\left(b+\sum_{i=1}^n a_i\mu_i,\quad \sum_{i=1}^n a_i^2\sigma_i^2\right)$$

2. 一维化二维：$X\sim N(\mu_1,\sigma_1^2)$，$Y\sim N(\mu_2,\sigma_2^2)$，且 X,Y 相互独立

$$\Longleftrightarrow (X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,0)$$

3. 二维化一维：若 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，则 X,Y 的任意线性组合

$aX+bY+c$ 服从一维正态分布 $N(\hat{\mu},\hat{\sigma}^2)$，其中参数

$$\hat{\mu}=E(aX+bY+c)=a\mu_1+b\mu_2+c$$

$$\begin{aligned}\hat{\sigma}^2&=D(aX+bY+c)=D(aX+bY)=\mathrm{Cov}(aX+bY,aX+bY)\\&=a^2\sigma_1^2+b^2\sigma_2^2+2\rho ab\sigma_1\sigma_2\end{aligned}$$

4. 二维协方差：若 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，则

$$\mathrm{Cov}(X,Y)=\rho_{XY}\cdot\sqrt{D(X)\cdot D(Y)}=\rho\sigma_1\sigma_2$$

$$E(XY)=\mathrm{Cov}(X,Y)+E(X)E(Y)=\rho\sigma_1\sigma_2+\mu_1\mu_2$$

4　随机变量数字特征

4.1　均值/数学期望

以下当所属求和式或积分式绝对收敛时，均值存在。

1. 一维随机变量：
   - 离散 X，分布列 $p_i=P(X=x_i)$，则 $E(X)=\sum\limits_{i=1}^{\infty}x_ip_i$
   - 连续 X，密度 $f(x)$，则 $E(X)=\int_{-\infty}^{+\infty}xf(x)\mathrm{d}x$
2. 随机变量函数 $Y=g(X)$：
   - 离散 X，分布列 $p_i=P(X=x_i)$，则 $E(Y)=E(g(X))=\sum\limits_{i=1}^{\infty}g(x_i)\cdot P(X=x_i)$
   - 连续 X，密度 $f(x)$，则 $E(Y)=E(g(X))=\int_{-\infty}^{+\infty}g(x)f(x)\mathrm{d}x$
3. 二维随机变量函数 $Z=g(X,Y)$：
   - 离散 (X,Y)，分布列 $p_{ij}=P(X=x_i,Y=y_j)$，则 $E(Z)=E(g(X,Y))=\sum\limits_{i=1}^{\infty}\sum\limits_{j=1}^{\infty}g(x_i,y_j)\cdot p_{ij}$
   - 连续 (X,Y)，密度 $f(x,y)$，则 $E(Z)=E(g(X,Y))=\iint_{\mathbb{R}^2}g(x,y)f(x,y)\mathrm{d}x\mathrm{d}y$

k 阶原点矩：$E(X^k)$

k 阶中心矩：$E\left\{[X-E(X)]^k\right\}$

4.2 方差

方差定义为二阶中心矩：$D(X) := E\left\{[X - E(X)]^2\right\}$

标准差：$\sqrt{D(X)}$

计算方差：$D(X) = E(X^2) - [E(X)]^2$

	均值	方差
常数	$E(c) = c$	$D(c) = 0$
数乘	$E(aX) = aE(X)$	$D(aX) = a^2 D(X)$
加常数	$E(X+b) = E(X) + b$	$D(X+b) = D(X)$
可加性	$E(X+Y) = E(X) + E(Y)$	【独立】$D(X+Y) = D(X) + D(Y)$
可乘性	【独立】$E(XY) = E(X)E(Y)$	—

表 2: 均值和方差的性质

4.3 协方差、相关系数

协方差：$\mathrm{Cov}(X,Y) = E(XY) - E(X)E(Y)$

- 自反性：$\mathrm{Cov}(X,X) = D(X)$
- 对称性：$\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$
- 数乘：$\mathrm{Cov}(aX,bY) = ab\mathrm{Cov}(X,Y)$
- 加常数：$\mathrm{Cov}(X+a, Y+b) = \mathrm{Cov}(X,Y)$
- 分配律：$\mathrm{Cov}(X_1+X_2, Y) = \mathrm{Cov}(X_1,Y) + \mathrm{Cov}(X_2,Y)$
- 独立：若 X,Y 独立，则 $\mathrm{Cov}(X,Y) = 0$
- 与方差的关系：

$$D(aX+bY) = \mathrm{Cov}(aX+bY, aX+bY) = a^2 D(X) + 2ab\mathrm{Cov}(X,Y) + b^2 D(Y)$$

相关系数：

$$\rho_{XY} := \frac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)} \cdot \sqrt{D(Y)}} \tag{8}$$

它总是 $[-1,1]$ 之间的实数。

- $\rho = 1$ 为正线性相关，即存在 $a > 0, b \in \mathbb{R}$ 使得 $P(Y = aX + b) = 1$
- $\rho = -1$ 为负线性相关，即存在 $a < 0, b \in \mathbb{R}$ 使得 $P(Y = aX + b) = 1$
- $\rho = 0$ 为互不相关，此时 $\mathrm{Cov}(X,Y) = 0$。相互独立 $\Longrightarrow$ 互不相关。

4.4 其他内容

随机变量标准化：X 为均值和方差都存在的随机变量，$Y=\dfrac{X-E(X)}{\sqrt{D(X)}}$，则 $E(Y)=0$，$D(Y)=1$。

切比雪夫不等式：对任意 $a>0$，

$$P\left(|X-E(X)|\geqslant a\right)\leqslant\frac{D(X)}{a^2}\tag{9}$$

推论：$D(X)=0\iff P(X=E(X))=1$。用于估计随机变量偏离中心的程度。

辛钦大数定律：设 $X_1,X_2,\ldots,X_n$ 相互独立同分布，均值为 μ，方差可以不存在，则 $\frac{1}{n}\sum_{i=1}^n X_i\xrightarrow{P}\mu$。样本均值依概率收敛到总体均值。

5 数理统计、区间估计、点估计

5.1 数理统计、抽样分布

总体：研究对象的全体。

样本：抽取的部分个体。

简单随机样本：相互独立且与总体 X 同分布的样本 $(X_1,X_2,\ldots,X_n)$。

统计量：关于样本的函数 $T(X_1,X_2,\ldots,X_n)$，不含未知参数。

名称	表达式
样本均值	$\bar{X}=\dfrac{1}{n}\sum\limits_{i=1}^n X_i$
样本方差	$S^2=\dfrac{1}{n-1}\sum\limits_{i=1}^n(X_i-\bar{X})^2$
样本标准差	$S=\sqrt{\dfrac{1}{n-1}\sum\limits_{i=1}^n(X_i-\bar{X})^2}$
样本 k 阶原点矩	$A_k=\dfrac{1}{n}\sum\limits_{i=1}^n X_i^k$
极大次序统计量	$X_{(n)}=\max\{X_1,X_2,\ldots,X_n\}$
极小次序统计量	$X_{(1)}=\min\{X_1,X_2,\ldots,X_n\}$

表 3: 常见的统计量

统计量的常用分布：标准正态分布、χ^2 分布、t 分布、F 分布。详见第二页。

单正态总体抽样分布：总体 $X \sim N(\mu,\sigma^2)$，样本 $(X_1,X_2,\ldots,X_n)$，样本均值 $\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$，样本方差 $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$，则有

1. $\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$
2. $\dfrac{n-1}{\sigma^2}S^2 \sim \chi^2(n-1)$，且 $\bar{X}$ 与 S^2 独立
3. $\dfrac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$

双正态总体抽样分布：X,Y 相互独立

总体	样本	样本均值	样本方差	样本容量
$X \sim N(\mu_1,\sigma_1^2)$	$(X_1,\ldots,X_n)$	$\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$	$S_1^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$	n
$Y \sim N(\mu_2,\sigma_2^2)$	$(Y_1,\ldots,Y_m)$	$\bar{Y}=\frac{1}{m}\sum_{i=1}^{m}Y_i$	$S_2^2=\frac{1}{m-1}\sum_{i=1}^{m}(Y_i-\bar{Y})^2$	m

1. $\dfrac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n}+\dfrac{\sigma_2^2}{m}}} \sim N(0,1)$
2. $\dfrac{S_1^2}{S_2^2}\cdot\dfrac{\sigma_2^2}{\sigma_1^2} \sim F(n-1,m-1)$
3. 当 $\sigma_1^2=\sigma_2^2$ 时有 $\dfrac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sqrt{\dfrac{(n-1)S_1^2+(m-1)S_2^2}{n+m-2}}\cdot\sqrt{\dfrac{1}{n}+\dfrac{1}{m}}} \sim t(n+m-2)$

上 α 分位点：满足 $P(Y>Y_\alpha)=\alpha$ 的分位点 Y_α。

构造上 α 分位点：

1. 大概率区间
   - $P(Y_{1-\frac{\alpha}{2}}<Y<Y_{\frac{\alpha}{2}})=1-\alpha$
   - $P(Y<Y_\alpha)=1-\alpha$
   - $P(Y>Y_{1-\alpha})=1-\alpha$
2. 小概率区间
   - $P(Y<Y_{1-\frac{\alpha}{2}}$ 或 $Y>Y_{\frac{\alpha}{2}})=\alpha$

- $P(Y > Y_\alpha) = \alpha$
- $P(Y < Y_{1-\alpha}) = \alpha$

常用的统计量上 α 分位点如表所示。

分布类型	Y_α	概率	特殊性质
$Z \sim N(0,1)$	z_α	$P(Z > z_\alpha) = \alpha$	$z_{1-\alpha} = -z_\alpha$
$t \sim t(n)$	$t_\alpha(n)$	$P(t > t_\alpha(n)) = \alpha$	$t_{1-\alpha}(n) = -t_\alpha(n)$
$\chi^2 \sim \chi^2(n)$	$\chi^2_\alpha(n)$	$P(\chi^2 > \chi^2_\alpha(n)) = \alpha$	无
$F \sim F(n,m)$	$F_\alpha(n,m)$	$P(F > F_\alpha(n,m)) = \alpha$	$F_{1-\alpha}(n,m) = \dfrac{1}{F_\alpha(m,n)}$

表 4: 常用的统计量上 α 分位点

另外，对于带有对称性的 $Z \sim N(0,1)$ 和 $t \sim t(n)$，还有：

$$P(z_{1-\frac{\alpha}{2}} < Z < z_{\frac{\alpha}{2}}) = 1-\alpha \iff P(|Z| < z_{\frac{\alpha}{2}}) = 1-\alpha$$
$$P(Z < z_{1-\frac{\alpha}{2}} \text{或} Z > z_{\frac{\alpha}{2}}) = \alpha \iff P(|Z| > z_{\frac{\alpha}{2}}) = \alpha$$

t 分布同理。

5.2 置信区间与假设检验

总体 X，分布 $F(x;\theta)$，其中 θ 待估，样本 $(X_1, X_2, \ldots, X_n)$。T_1, T_2 是关于样本的统计量。

- $P(T_1 \leqslant \theta \leqslant T_2) = 1-\alpha$，置信区间为 $[T_1, T_2]$
- $P(T_1 \leqslant \theta) = 1-\alpha$，置信下限为 T_1
- $P(\theta \leqslant T_2) = 1-\alpha$，置信上限为 T_2

其中置信度是 $1-\alpha$。

求置信区间的解法：

1. 构造样本函数（单参数）
2. 构造对应的大概率事件
3. 解出置信区间
4. 代值计算

假设检验：认为「一次试验中小概率事件不会发生」。若发生了，就拒绝原假设 H_0；否则接受 H_0。

两种错误：一为弃真，概率为 α；二为采伪，概率为 β。α,β 相互制约，要同时降低它们只能增加样本容量。这里只考虑弃真错误，减小 α，称为「显著性检验」。α 称为显著性水平。

假设检验方法：

1. 提出假设
2. 选取检验统计量
3. 构造对应的小概率事件
4. 代值计算，判定统计量是否落入拒绝域

$$
\begin{array}{c|c|c}
H_0:\theta=\theta_0 & H_0:\theta=\theta_0 & H_0:\theta=\theta_0 \\
H_1:\theta\neq\theta_0 & H_1:\theta<\theta_0 & H_1:\theta>\theta_0 \\
P(Y<Y_{1-\frac{\alpha}{2}}\text{或}Y>Y_{\frac{\alpha}{2}})=\alpha & P(Y<Y_{1-\alpha})=\alpha & P(Y>Y_{\alpha})=\alpha
\end{array}
$$

5.3 点估计与优良性

总体 X，分布 $F(x;\theta)$，其中 θ 待估，则 $E(X)=h(\theta)$。

矩估计：令样本原点矩等于总体原点矩，如一阶 $\frac{1}{n}\sum_{i=1}^{n}X_i=E(X)=h(\theta)$。解出来 $\hat{\theta}=\hat{\theta}(X_1,X_2,\ldots,X_n)$ 是 θ 的矩估计量，代样本观测值得到的 $\hat{\theta}=\hat{\theta}(x_1,x_2,\ldots,x_n)$ 则是 θ 的矩估计值。（依据：辛钦大数定律）

$$\frac{1}{n}\sum_{i=1}^{n}X_i^k \xrightarrow[\text{依概率收敛}]{P} E(X^k)$$

如果是多参数（m 个），则让一阶、二阶、⋯⋯、m 阶样本原点矩等于总体原点矩，解方程组。

极大似然估计：样本似然函数 $L(\theta)=\prod\limits_{i=1}^{n}p(x_i;\theta)$ 的最大值点 $\hat{\theta}$ 就是极大似然估计量/估计值。

似然函数中的 $p(x_i,\theta)$ 为分布列或者密度函数：

- 对离散型 X，$L(\theta)=\prod\limits_{i=1}^{n}p(x_i;\theta)=\prod\limits_{i=1}^{n}P(X=x_i)$
- 对连续型 X，$L(\theta)=\prod\limits_{i=1}^{n}p(x_i;\theta)=\prod\limits_{i=1}^{n}f(x_i;\theta)$

为了确定最大值点 $\hat{\theta}$，一般采用取对数法或者利用单调性等。

无偏性：若总体参数 θ 的估计量 $\hat{\theta}$ 满足 $E(\hat{\theta})=\theta$，则称 $\hat{\theta}$ 是无偏估计量。

有效性：在参数 θ 的所有无偏估计量中，称方差小者更有效。

6 可能用到的工数知识

1. 求多元函数极值——拉格朗日乘数法。

   在约束条件 $\varphi(x,y,z)=0$ 下，求函数 $f(x,y,z)$ 的极值。令 $L(x,y,z;\lambda)=f(x,y,z)+\lambda\cdot\varphi(x,y,z)$，同时解四元方程组

$$\frac{\partial L}{\partial x}=\frac{\partial L}{\partial y}=\frac{\partial L}{\partial z}=\frac{\partial L}{\partial \lambda}=0 \tag{10}$$

   解得的 (x_0,y_0,z_0) 就是函数 $f(x,y,z)$ 的可能极值点。

2. 直角坐标与极坐标积分互化。

$$\iint\limits_{D} f(x,y)\mathrm{d}x\mathrm{d}y=\iint\limits_{D_{r\theta}} f(r\cos\theta,r\sin\theta)\cdot r\mathrm{d}r\mathrm{d}\theta \tag{11}$$

3. 分部积分。

$$\int f'(x)g(x)\mathrm{d}x=f(x)g(x)-\int g'(x)f(x)\mathrm{d}x \tag{12}$$

4. 变限积分求导。当 $f(x)$ 连续时，

$$\left(\int_{\alpha(x)}^{\beta(x)} f(t)\mathrm{d}t\right)'=f(\beta(x))\cdot\beta'(x)-f(\alpha(x))\cdot\alpha'(x) \tag{13}$$