01 连续型数学期望

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

$E(Y)=E(g(x))=\int_{-\infty}^{+\infty} g(x) f(x)\,\mathrm{d}x$

$E(Z)=E(h(X,Y))=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} h(x,y) f(x,y)\,\mathrm{d}x\,\mathrm{d}y$

02 方差的定义

$D(X)=\mathrm{Var}(X)=E((X-E(X))^2)=E(X^2)-(E(X))^2$。

根据定义，方差实际上是随机变量 X 的函数$g(X)=(X-E(X))^2$的数学期望。

03 切比雪夫不等式

$P(|X-\mu|\geq\varepsilon)\leq\frac{\sigma^2}{\varepsilon^2}$

04 协方差的定义

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

相关系数：$\rho_{XY}=\frac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)\cdot D(Y)}}\in[-1,1]$。

$D(X+Y)=D(X)+D(Y)+2\ \mathrm{Cov}(X,Y)$

$=D(X)+D(Y)$(特别地，当 X 和 Y 独立时)

05 数学期望、方差、协方差的性质

$E(CX)=CE(X)$，$E(aX+bY)=aE(X)+bE(Y)$

$D(CX)=C^2D(X)$，$D(X\pm Y)=D(X)+D(Y)$，$D(X+C)=D(X)$。

$\mathrm{Cov}(X,Y)=\mathrm{Cov}(Y,X),\mathrm{Cov}(X,X)=D(x),\mathrm{Cov}(X,a)=0,\mathrm{Cov}(aX+bY)=ab\ \mathrm{Cov}(X,Y),\mathrm{Cov}(X_1+X_2,Y)=\mathrm{Cov}(X_1,Y)+\mathrm{Cov}(X_2,Y)$。

06 特殊分布的期望和方差

1. 泊松分布$X\sim P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda}$：$E(X)=\lambda,D(x)=\lambda$。
2. 指数分布$f(x)=\frac{1}{\theta}e^{-\frac{x}{\theta}}(\text{when } x>0)$：$E(X)=\theta,D(X)=\theta^2$。
3. 均匀分布$f(x)=\frac{1}{b-a}(\text{when } a<x<b)$：$E(X)=\frac{a+b}{2},D(X)=\frac{(b-a)^2}{12}$。

07 矩

$E(X^k)$：k阶（原点）矩。

$E(X^kY^l)$：$k+l$阶混合矩。

$E\{(X-E(X))^k\}$：k阶中心矩。

$E\{(X-E(X))^k(Y-E(Y))^l\}$：$k+l$阶混合中心矩。

08 大数定理

- 切比雪夫大数定理：随机变量相互独立，且具有相同的数学期望和方差：$E(X_k)=\mu,D(X_k)=\sigma^2$，作前 n 个随机变量的算术平均$\overline{X}=\frac{\sum_{k=1}^{n}}{n}$，则有：

$\lim_{n\to\infty}P\{|\overline{X}-\mu|<\varepsilon\}=1$，即$\overline{X}$依概率收敛于$\mu$。

- 伯努利大数定理：n_A是 n 次独立重复试验中事件 A 发生的次数，p是事件 A 在每次试验中发生的概率 $\Rightarrow\frac{n_A}{n}$依概率收敛于μ。
- 辛钦大数定理：随机变量独立同分布，且具有相同的数学期望$E(X_k)=\mu\Rightarrow\overline{X}$依概率收敛于$\mu$

09 独立同分布的中心极限定理

随机变量独立同分布，且具有相同的数学期望和方差：$E(X_k)=\mu,D(X_k)=\sigma^2$，则随机变量之和的标准化变量：$Y_n=\frac{\sum_{k=1}^{n}-n\mu}{\sqrt{n}\sigma}$的分布函数对于任意$x$满足$\lim_{n\to\infty}P\left\{\frac{\sum_{k=1}^{n}-n\mu}{\sqrt{n}\sigma}\right\}\leq x=\Phi(x)$，即收敛于标准正态分布的分布函数。

10 简单随机抽样

$X_1,X_2,...,X_n$是F的一个样本，则它们的联合分布函数为$F^*(x_1,x_2,...,x_n)=\prod_{i=1}^{n}F(x_i)$，联合概率密度函数（如果存在）为$f^*(x_1,x_2,...,x_n)=\prod_{i=1}^{n}f(x_i)$。

11 常用统计量的定义

- 样本均值：$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$。

对于正态总体的样本，$\overline{X}\sim N\left(\mu,\frac{\sigma^2}{n}\right)$。

- 样本方差：$S^2=\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2=\frac{1}{n-1}\left(\sum_{i=1}^{n}X_i^2-n\overline{X}^2\right)$。

对于正态总体的样本，有$\frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)$，$\overline{X}$与$S^2$独立。

对于正态总体的样本，$\frac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}}\sim t(n-1)$。

12 χ^2分布

- 定义：$\chi^2=X_1^2+X_2^2+...+X_k^2\sim\chi^2(n)$，其中$X_i$是来自总体$N(0,1)$的独立随机变量。
- 期望和方差：$E(X)=n,D(X)=2n$。
- 性质：有$\chi_1^2\sim\chi^2(n_1),\chi_2^2\sim\chi^2(n_2)$，则$\chi_1^2+\chi_2^2=\chi^2(n_1+n_2)$。

13 t分布

- 定义：设$X\sim N(0,1),Y\sim\chi^2(n)$，且$X$和$Y$相互独立，则$t=\frac{X}{\sqrt{\frac{Y}{N}}}\sim t(n)$。
- 性质：图形关于$t=0$对称，在$t=0$处取到最大值，在 n 足够大时近似于标准正态分布。

14 F分布

- 定义：设$X \sim \chi^2(n_1), Y \sim \chi^2(n_2)$，则$F = \frac{\frac{X}{n_1}}{\frac{Y}{n_2}} \sim F(n_1, n_2)$。
- 性质：若$F \sim F(n_1, n_2)$，则$\frac{1}{F} \sim F(n_2, n_1)$。

15 点估计的矩估计法

用样本矩的连续函数来估计总体矩的连续函数。

- 总体中只有一个待估参数，用样本的一阶矩来估计总体的一阶矩:

$$\begin{cases} A_1 = \frac{1}{n}\sum_{i=1}^{n} X_i \\ \mu_1 = E(X) \end{cases} \Rightarrow A_1 = \mu_1 \Rightarrow \text{解出一个未知参数}$$

- 总体中有两个待估参数，用样本的前二阶矩来估计总体的前二阶矩:

$$\begin{cases} A_1 = \frac{1}{n}\sum_{i=1}^{n} X_i \\ A_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 \\ \mu_1 = E(X) \\ \mu_2 = E(X^2) = D(X) + (E(X))^2 \end{cases} \Rightarrow \begin{cases} A_1 = \mu_1 \\ A_2 = \mu_2 \end{cases} \Rightarrow \text{解出两个未知参数}$$

16 假设检验

1. 提出假设H_0和备择假设H_1;
2. 构造一个统计量，满足假设H_0且不含任何未知参数;
3. 给定假设H_0成立的概率为$1-\alpha$（显著性水平），这个统计量的取值应当不落入一定的范围W（拒绝域）;
4. 根据样本的观察值，计算统计量的观察值u_0，若t_0落入拒绝域，则拒绝假设H_0。

针对 正态总体的参数假设，可以按照下表 进行假设检验：

$X \sim N(\mu, \sigma^2)$

统计假设		条件	构造的统计量及其分布	拒绝域
H_0	H_1			
$\mu = \mu_0$	$\mu \neq \mu_0$	σ^2 已知	$U = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$	$\lvert u\rvert \geq u_{\alpha/2}$
$\mu \leq \mu_0$	$\mu > \mu_0$			$u \geq u_\alpha$
$\mu \geq \mu_0$	$\mu < \mu_0$			$u \leq -u_\alpha$
$\mu = \mu_0$	$\mu \neq \mu_0$	σ^2 未知	$T = \frac{\bar{X} - \mu_0}{S_n^*/\sqrt{n}} \sim t(n-1)$	$\lvert t\rvert \geq t_{\alpha/2}(n-1)$
$\mu \leq \mu_0$	$\mu > \mu_0$			$t \geq t_\alpha(n-1)$
$\mu \geq \mu_0$	$\mu < \mu_0$			$t \leq -t_\alpha(n-1)$
$\sigma^2 = \sigma_0^2$	$\sigma^2 \neq \sigma_0^2$	μ 未知	$\chi^2 = \frac{n-1}{\sigma_0^2} S_n^{*2} \sim \chi^2(n-1)$	$\chi^2 \geq \chi^2_{\alpha/2}(n-1)$ 或 $\chi^2 \leq \chi^2_{1-\alpha/2}(n-1)$
$\sigma^2 \leq \sigma_0^2$	$\sigma^2 > \sigma_0^2$			$\chi^2 \geq \chi^2_\alpha(n-1)$
$\sigma^2 \geq \sigma_0^2$	$\sigma^2 < \sigma_0^2$			$\chi^2 \leq \chi^2_{1-\alpha}(n-1)$