

(空)	1	6.67%
本题有效填写人次	15	



第6题：其它关于「概率论与数理统计（智）」的建议 [填空题]

词频分析

详情

答卷数据有更新，您可以获取最新分析结果

提交更新

讲细

第5题：如你认为必要性>3,概率论中你认为问题较大的部分 [多选题]

选项	小计	比例
期中前内容（基本定义、数字特征、大数定律）	7	46.67%
统计基本概念（总体样本、统计量）	6	40%
正态总体重要分布	6	40%
参数估计	12	80%
假设检验	14	93.33%
随机过程基本	12	80%
平稳过程	10	66.67%
其它 [详情]	0	0%
(空)	1	6.67%
本题有效填写人次	15	

目录

- Chap(1-5)过去的章节
- Chap(6-8)几个重要的分布
- 重要关系
- 上下分位数
- 样本统计量的无偏性
- 区间估计
- 矩估计、最大似然估计
- 假设检验
- Chap(11-12)随机过程
- 傅里叶变换的理解
- 怎么不查表求 $S(w)$
- 各态历经性

内容比较基础，以过知识点
(唤醒沉睡的记忆) 为主

! 凌乱草稿预警

! 错误预警

! 主讲者把自己讲晕了预警

怎么求概率密度函数 $f_X(x)$ (或者是分布律)

(当然是连续的)

1. 分布律：瞪眼法。 概率密度： $F'(x) = f(x)$
2. 没有 $F(x)$ ，把它算出来再求导（书本p44 例2.15）
3. 算不出来，看看有无其他方法
 1. 书本p43随机变量的函数：已知 $f_X(x)$ ， $y=g(x)$ 且 $g(x)$ 单调，那么 $f_Y(y)$ 公式；
 2. 部分分布的可加性（正态，泊松， Γ ， χ^2 ）
 3. 二维随机变量 $f(x,y)$ ：高数里学过的积分变换（注意值域和独立性）

一般地，有以一定义。
定理 2.1 设随机变量 X 具有概率密度 $f_X(x)$ ，若 $g(x)$ 为 $(-\infty, +\infty)$ (即实数域 $\mathbf{R}$) 上的可导函数，并且，恒有 $g'(x) > 0$ (或恒有 $g'(x) < 0$)。记 $h(y)$ 为 $g(x)$ 的反函数，则 $Y=g(X)$ 是连续型随机变量，其概率密度为

$$f_Y(y) = \begin{cases} f_X(h(y)) |h'(y)|, & y \in g(\mathbf{R}), \\ 0, & \text{其他} \end{cases}$$

其中 $g(\mathbf{R}) = \{g(x) \mid x \in (-\infty, +\infty)\}$ 为 $g(x)$ 的值域。

$g_2(X, Y)$ 是 (X, Y) 到 (U, V) 的一一对应变换，即存在逆变换 $X=h_1(U, V)$, $Y=h_2(U, V)$ ，并且 g_1, g_2, h_1, h_2 都存在连续的一阶偏导数，则 (U, V) 仍为连续型随机变量，其概率密度为

$$f_{U,V}(u, v) = \begin{cases} f(h_1(u, v), h_2(u, v)) |J|, & (u, v) \in \mathcal{B}, \\ 0, & \text{其他} \end{cases}$$

其中 $\mathcal{B}$ 为向量值函数 $(u, v) = (g_1(x, y), g_2(x, y))$ 的值域。

记忆唤醒:

即 X 与 Y 相互独立。

例 3.9 设二维随机变量 (X, Y) 的联合概率密度为


$$f(x, y) = \begin{cases} 3x, & 0 < x < 1, 0 < y < x \\ 0, & \text{其他} \end{cases}$$

求 $Z = X - Y$ 的概率密度。

解 先求 Z 的分布函数

例 3.12 在区间 $[0, 1]$ 上随机地投掷两点, 试求这两点间距离的概率密度。

答案: $\frac{3}{2}(1-z^2) \ 0 \leq z \leq 1; \ 2(1-z) \ 0 \leq z < 1$

练习： $X_1, X_2, \dots, X_n$ 是独立同分布的随机变量， X_M 为 $\max\{X_1, X_2, \dots, X_n\}$, 求 X_M 的概率密度函数。

简单一点： 均匀分布 $U(0,1)$

Min?

离散情况？

有了概率密度函数可以干什么

$$E[g(X)]$$

$$E[X]$$

$$E[X^k] \quad \rightarrow \quad D(x)$$

二元： $E[g(X,Y)]$

求边缘概率分布 – 判断是否独立

求条件概率密度

求协方差

独立与相关

独立的原始定义：

二维随机向量 (X, Y) , 满足 $F(x,y)=F_x(x) F_y(y)$, 其中 $F_x(x) F_y(y)$ 是边缘分布函数。

离散： $P\{X=x_i, Y=y_j\} = P\{X=x_i\}P\{Y=y_j\}$

连续： $f(x,y) = f_x(x)f_y(y)$

若独立， 则有 $E(XY)=E(X)E(Y)$, 于是就有 $Cov(X,Y)=0$ 即独立 $\rightarrow$ 不相关
但是不相关不能推出独立（正态分布和二元分布除外）

两个与独立有关的定理

定理 3.1 设 $(X_1, X_2, \dots, X_m)$ 与 $(Y_1, Y_2, \dots, Y_n)$ 相互独立, 若 $h(x_1, x_2, \dots, x_m)$, $g(y_1, y_2, \dots, y_n)$ 是连续函数, 则 $h(X_1, X_2, \dots, X_m)$ 与 $g(Y_1, Y_2, \dots, Y_n)$ 也相互独立.

定理 3.2 若随机变量 $X_1, X_2, \dots, X_n$ 相互独立, 把它们分为不相交的 k 组, 每组中的所有变量由一个连续函数复合而生成一个新的随机变量, 则这 k 个随机变量仍相互独立.

几个重要的分布

Γ分布

定义6.2: 如果随机变量 X 具有概率密度

$$f(x, \alpha, \lambda) = \begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

其中 $\alpha > 0, \lambda > 0$ 为常数, 则称 X 服从参数为 α, λ 的 Γ 分布, 记为 $X \sim \Gamma(\alpha, \lambda)$.

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

$$\Gamma(\alpha + 1) = \alpha \Gamma(\alpha)$$

指数分布为 $\alpha=1$ 的 Γ 分布

性质1: 如果随机变量 $X \sim \Gamma(\alpha, \lambda)$, 那么 $E(X) = \frac{\alpha}{\lambda}$,

$$D(X) = \frac{\alpha}{\lambda^2}.$$

性质2: 如果随机变量 $X_1 \sim \Gamma(\alpha_1, \lambda), X_2 \sim \Gamma(\alpha_2, \lambda)$, 且

X_1 和 X_2 相互独立, 则 $X_1 + X_2 \sim \Gamma(\alpha_1 + \alpha_2, \lambda)$ **可加 (对 α 可加)**

几个重要的分布

2. χ^2 分布

在 Γ 分布中取 $\alpha = \frac{n}{2}, \lambda = \frac{1}{2}$, Γ 分布就是自由度为 n 的卡方分布

定义6.3: 若随机变量 Z 具有概率密度

$$\chi^2(x, n) = \begin{cases} \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, & x > 0, \\ 0, & x \leq 0 \end{cases}$$

则称 Z 服从自由度为 n 的 χ^2 分布, 记作 $Z \sim \chi^2(n)$,

性质1: 如果随机变量 $Z \sim \chi^2(n)$, 那么 $E(Z) = n$,

$D(Z) = 2n$. **均值和方差**

性质2: 如果随机变量 $Z_1, Z_2, \dots, Z_m$ 相互独立, 且

$Z_i \sim \chi^2(n_i) (i=1, 2, \dots, m)$, 则 $\sum_{i=1}^m Z_i \sim \chi^2(n_1 + n_2 + \dots + n_m)$.
可加

可表示为 n 个正态变量的平方和

定理6.4 设随机变量 $X_1, X_2, \dots, X_n$, 且都服从标准正态分布 $N(0,1)$, 则随机变量 $\chi^2 = \sum_{i=1}^n X_i^2$ 服从自由度为 n 的卡方分布, 即 $\chi^2 \sim \chi^2(n)$.

$(N(0,1))^2 \sim \chi^2(1)$

几个重要的分布

3. t 分布

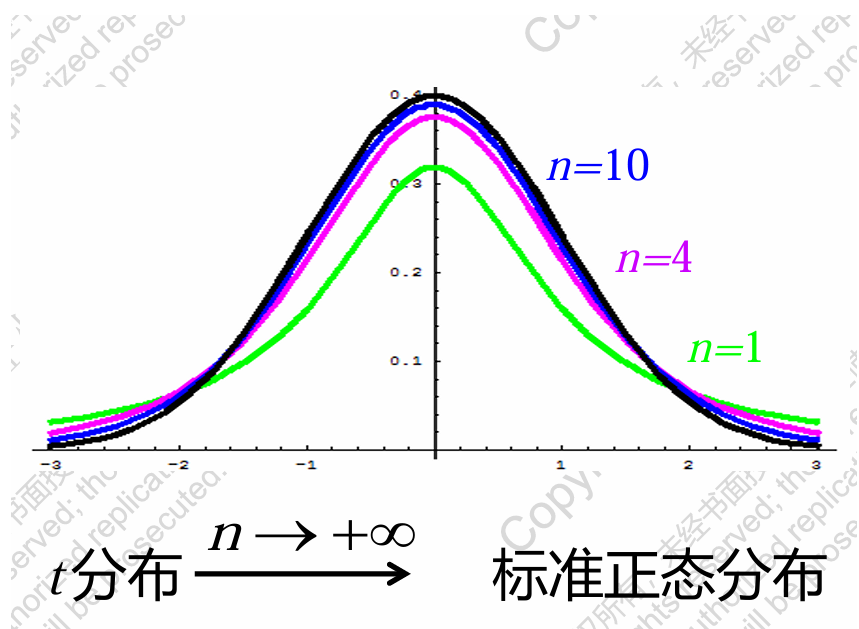
定义6.4: 若随机变量 T 具有概率密度函数

$$t(x; n) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi n} \Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}, \quad -\infty < x < +\infty$$

则称 T 服从自由度为 n 的 t 分布, 记作 $T \sim t(n)$.

t 分布又称为学生氏(Student)分布

长得像正态分布, 对称



几个重要的分布

4. F 分布

定义6.5: 若随机变量 F 具有概率密度服从

$$f(x, n, m) = \begin{cases} \frac{\Gamma\left(\frac{n+m}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{m}{2}\right)} \left(\frac{n}{m}\right)^{\frac{n}{2}} x^{\frac{n}{2}-1} \left(1 + \frac{n}{m}x\right)^{-\frac{n+m}{2}}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

则称 F 服从自由度为 (n, m) 的 F 分布,

记作 $F \sim F(n, m)$

重要关系

定理6.5 若 $X \sim N(0, 1)$, $Y \sim \chi^2(n)$, 并且 X 与 Y 相互独立,

则
$$T = \frac{X}{\sqrt{Y/n}} \sim t(n)$$

定理6.6 若 $X \sim \chi^2(n)$, $Y \sim \chi^2(m)$, 且 X 与 Y 相互独立, 则

$$F = \frac{X/n}{Y/m} \sim F(n, m)$$

推论 若 $F \sim F(n, m)$, 则

$$\frac{1}{F} \sim F(m, n)$$

习题里的结论: 总体 $X \sim \text{Exp}(\lambda)$, 则 $2n\lambda \bar{X} \sim \chi^2(2n)$

重要关系

服从自由度为 (n_1, n_2) 的 F 分布

25. 设总体 X 服从参数为 λ 的指数分布, $(X_1, X_2, \dots, X_n)$ 是来自总体 X 的样本, 试证明 $Y =$

$2\lambda \sum_{i=1}^n X_i$ 服从自由度为 $2n$ 的 χ^2 分布.

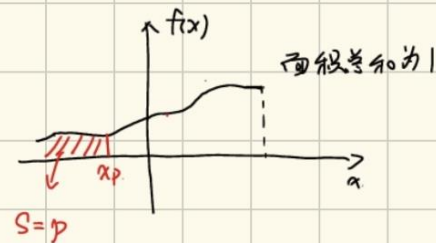
27. 设 $(X_1, X_2, \dots, X_n)$ 为来自总体 $U(0, 1)$ 的样本, 试证明 $Y = -2 \sum_{i=1}^n \ln(X_i)$ 服从自由度为 $2n$

的 χ^2 分布.

上下分位数

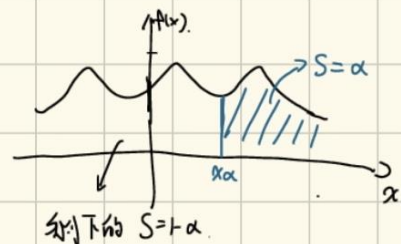
下侧 p 分位数

$$P\{X \leq x_p\} = F(x_p) = p$$



上侧 α 分位数

$$P\{X > x_\alpha\} = 1 - F(x_\alpha) = \alpha$$



因此知道 $f(x)$

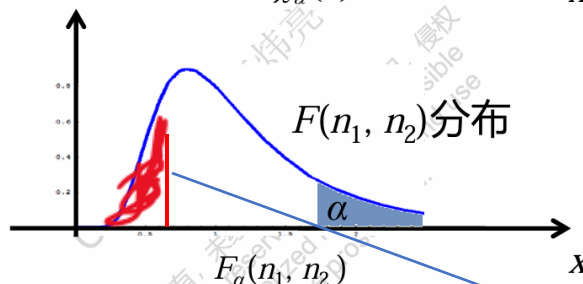
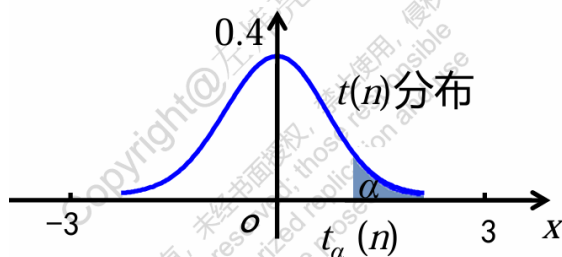
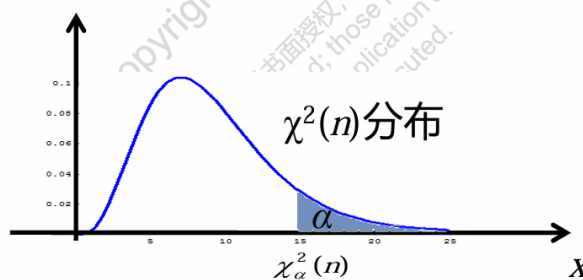
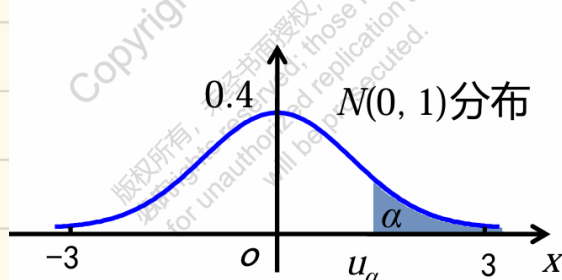
可以通过积分求出分位数

$$F(x_p) = \int_{-\infty}^{x_p} f(x) dx = p \quad \text{解 } x_p = x_\alpha$$

$$F(x_\alpha) = \int_{-\infty}^{x_\alpha} f(x) dx = 1 - \alpha$$

在后续的参数估计、假设检验中，应用 α 分位数比较多

上侧 α 分位数图示



$F_{1-\alpha}(n_1, n_2)$

性质

1. 对称

$$u_{1-\alpha} = -u_\alpha, \quad t_{1-\alpha}(n) = -t_\alpha(n)$$

2. F分布

$$F_\alpha(n, m) = \frac{1}{F_{1-\alpha}(m, n)}$$

证明

$X \sim F(n, m)$, 设 $F_\alpha(n, m)$ 为 $F(n, m)$ 的上侧 α 分位数.

$$P\{X > F_\alpha(n, m)\} = \alpha$$
$$P\left\{\frac{1}{X} < \frac{1}{F_\alpha(n, m)}\right\} = \alpha \rightarrow P\left\{\frac{1}{X} > \frac{1}{F_\alpha(n, m)}\right\} = 1 - \alpha.$$

$X \sim F(n, m)$, 设 $F_\alpha(n, m)$ 为 $F(n, m)$ 的上侧 α 分位数.

$$P\{X > F_\alpha(n, m)\} = \alpha$$

$$P\left\{\frac{1}{X} < \frac{1}{F_\alpha(n, m)}\right\} = \alpha \rightarrow P\left\{\frac{1}{X} > \frac{1}{F_\alpha(n, m)}\right\} = 1 - \alpha.$$

又因为 $\frac{1}{F(n, m)} \sim F(m, n)$ 即 $Y = \frac{1}{X} \sim F(m, n)$

则 $\frac{1}{F_\alpha(n, m)}$ 为 Y 的上侧 $1 - \alpha$ 分位数.

$$\therefore \frac{1}{F_\alpha(n, m)} = F_{1-\alpha}(m, n).$$

样本统计量

设 $(X_1, X_2, \dots, X_n)$ 是来自总体 X 的样本, $(x_1, x_2, \dots, x_n)$ 是这一样本的观测值.

$$(1) \quad \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i = \frac{1}{n} (X_1 + X_2 + \dots + X_n) \quad \text{--- 样本均值}$$

$$(2) \quad S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \quad \text{--- 样本方差}$$

$$(3) \quad S = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2} \quad \text{--- 样本标准差}$$

$$(4) \quad A_k = \frac{1}{n} \sum_{i=1}^n X_i^k \quad \text{--- 样本的 } k \text{ 阶原点矩}$$

$$(5) \quad B_k = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^k \quad \text{--- 样本的 } k \text{ 阶中心矩}$$

样本统计量的无偏性

回顾：无偏估计量（课本p156-161）

定义1 设 $\hat{\theta}$ 是未知参数 θ 的估计量，如果任给 $\theta \in \Theta$ ，

都有 $E(\hat{\theta}) = \theta$ ，则称 $\hat{\theta}$ 是 θ 的**无偏估计量**

课本结论：不论样本服从什么分布，样本的均值、方差、k阶原点矩都是总体对应统计量的无偏估计量（证明在p129-131）

但是对两者进行非线性变换，结果却不一定还是无偏估计（例7.19）

其他的评选标准：有效性（方差） 相合性（依概率收敛）

渐近无偏估计量

好像不是重点？有好心人去问问老师吗

正态总体：样本统计量的分布

定理6.7 设 $X_1, X_2, \dots, X_n$ 是来自正态总体 $N(\mu, \sigma^2)$ 的样本, $\bar{X}$ 为样本均值, S^2 为样本方差, 则

1) $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

2) $\frac{(n-1)S^2}{\sigma^2} = \sum_{i=1}^n \left(\frac{X_i - \bar{X}}{\sigma}\right)^2 \sim \chi^2(n-1)$

$$\sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma}\right)^2 \sim \chi^2(n)$$

3) $\bar{X}$ 与 S^2 相互独立.

定理6.10 设 $X_1, X_2, \dots, X_n$ 与 $Y_1, Y_2, \dots, Y_m$ 分别为来自正态总体 $N(\mu_1, \sigma_1^2)$, $N(\mu_2, \sigma_2^2)$ 的样本, 且两组样本相互独立, S_1 和 S_2 如定理6.9定义, 则

$$F = \frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2} \sim F(n-1, m-1)$$

定理6.8 设 $X_1, \dots, X_n$ 是来自正态总体 $N(\mu, \sigma^2)$ 的样本, $\bar{X}$ 为样本均值, S^2 为样本方差, 则

$$T = \frac{\sqrt{n}(\bar{X} - \mu)}{S} \sim T(n-1)$$

需要两个正态总体方差相同

定理6.9 设 $X_1, X_2, \dots, X_n$ 与 $Y_1, Y_2, \dots, Y_m$ 分别为来自正态总体 $N(\mu_1, \sigma^2)$, $N(\mu_2, \sigma^2)$ 的样本且两组样本相互独立, 记

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

$$\bar{Y} = \frac{1}{m} \sum_{j=1}^m Y_j$$

$$S_1^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

$$S_2^2 = \frac{1}{m-1} \sum_{j=1}^m (Y_j - \bar{Y})^2$$

则

$$T = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{1}{n} + \frac{1}{m}} \sqrt{\frac{(n-1)S_1^2 + (m-1)S_2^2}{n+m-2}}} \sim t(n+m-2)$$

记成 S_w

区间估计

区间估计

样本 + 样本统计量

($\bar{x}$, S^2 ...)

↓
样本观测值 $\bar{x}$, S^2

理论: 样本统计量应服从一定分布

样本服从另一个分布

↓ 可能已知

(带有同样的未知参数)

D. P. σ^2

(课本上以正态分布为主)

但也出现其他问题或其他)

方法: 用(样本统计量 + 未知参数)服从的已知分布

($t(n)$, $\chi^2(n)$, $N(\mu, \sigma^2)$,

$F(n_1, n_2)$, ...)

代入样本观测值

以及该分布的在 $1-\alpha$ 区间的上下限

解不等式, 求出未知参数的范围 (置信度 $1-\alpha$)

(1) 精度: $\hat{\theta}_2 - \hat{\theta}_1$ 越小越好; 一般是让 $E(\hat{\theta}_2 - \hat{\theta}_1)$ 越小越好

(2) 置信度: $P\{\hat{\theta}_1 < \theta < \hat{\theta}_2\}$ 越大越好 即选择的 α 更小

区间估计常见情形

正态总体

(1) σ^2 为已知时, 均值 μ 的置信区间

σ^2 为已知时, 因为 $\bar{X}$ 是 μ 的无偏估计, 且 $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$

对于给定的 $\alpha (0 < \alpha < 1)$, 由前面的

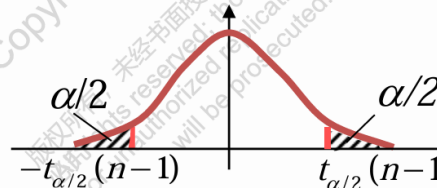
$$P\left\{\left|\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right| < z_{\alpha/2}\right\} = 1-\alpha \quad \text{即} \quad P\left\{\bar{X}-\frac{\sigma}{\sqrt{n}}z_{\alpha/2} < \mu < \bar{X}+\frac{\sigma}{\sqrt{n}}z_{\alpha/2}\right\} = 1-\alpha$$

故 μ 的置信度水平为 $(1-\alpha)$ 的置信区间

$$\left(\bar{X}-\frac{\sigma}{\sqrt{n}}z_{\alpha/2}, \bar{X}+\frac{\sigma}{\sqrt{n}}z_{\alpha/2}\right) \quad \text{或} \quad \left(\bar{X} \pm \frac{\sigma}{\sqrt{n}}z_{\alpha/2}\right)$$

(2) σ^2 未知, 求 μ 的置信区间

σ^2 为未知时, 因为 S^2 是 σ^2 的无偏估计量, 所以用 S 替换(1)中的 σ , 则有

$$\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$$


$$P\left\{-t_{\alpha/2}(n-1) < \frac{\bar{X}-\mu}{S/\sqrt{n}} < t_{\alpha/2}(n-1)\right\} = 1-\alpha$$

$$\text{有} \quad P\left\{\bar{X}-\frac{S}{\sqrt{n}}t_{\alpha/2}(n-1) < \mu < \bar{X}+\frac{S}{\sqrt{n}}t_{\alpha/2}(n-1)\right\} = 1-\alpha$$

求得 μ 的置信水平为 $(1-\alpha)$ 的置信区间 $\left(\bar{X} \pm \frac{S}{\sqrt{n}}t_{\alpha/2}(n-1)\right)$

区间估计常见情形

正态总体

(3) 方差 σ^2 的置信区间 (μ 未知)

σ^2 的无偏估计量为 S^2 ，当 $1-\alpha$ 给定后，

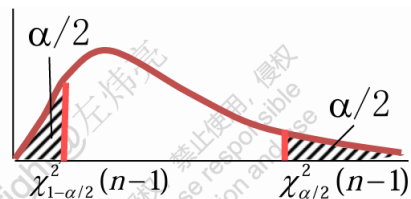
因为 $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$

故 $P\left\{\chi^2_{1-\alpha/2}(n-1) < \frac{(n-1)S^2}{\sigma^2} < \chi^2_{\alpha/2}(n-1)\right\} = 1-\alpha$

即 $P\left\{\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\right\} = 1-\alpha$

得到方差 σ^2 和标准差 σ 的一个置信度为 $1-\alpha$ 的置信区间分别为：

$$\left(\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}, \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\right) \text{ 和 } \left(\frac{\sqrt{n-1}S}{\sqrt{\chi^2_{\alpha/2}(n-1)}}, \frac{\sqrt{n-1}S}{\sqrt{\chi^2_{1-\alpha/2}(n-1)}}\right)$$



(1) 两个总体均值差 $\mu_1 - \mu_2$ 的置信区间 (置信度为 $1-\alpha$)

◆ σ_1^2, σ_2^2 均为已知: 因 $\bar{X} - \bar{Y}$ 为 $\mu_1 - \mu_2$ 的无偏估计量, 而

$$\bar{X} - \bar{Y} \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right) \text{ 故 } \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0, 1)$$

即得 $\mu_1 - \mu_2$ 的 $(1-\alpha)$ 置信区间

$$\left(\bar{X} - \bar{Y} \pm z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right)$$

◆ $\sigma_1^2 = \sigma_2^2 = \sigma^2$, 但 σ^2 为未知

由定理6.9

$$T = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2)$$

其中

$$S_w^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}, \quad S_w = \sqrt{S_w^2}$$

从而可得 $\mu_1 - \mu_2$ 的一个置信度为 $1-\alpha$ 的置信区间为

$$\left(\bar{X} - \bar{Y} \pm t_{\alpha/2}(n_1 + n_2 - 2) S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right)$$

两个正态总体

区间估计常见情形

两个正态总体

(2) 两个总体方差比 σ_1^2/σ_2^2 的置信区间

仅讨论总体均值 μ_1, μ_2 为未知的情况

由于 $\frac{S_1^2 / S_2^2}{\sigma_1^2 / \sigma_2^2} \sim F(n_1 - 1, n_2 - 1)$

$$P\left\{F_{1-\alpha/2}(n_1-1, n_2-1) < \frac{S_1^2 / S_2^2}{\sigma_1^2 / \sigma_2^2} < F_{\alpha/2}(n_1-1, n_2-1)\right\} = 1 - \alpha$$

$$\text{即 } P\left\{\frac{S_1^2}{S_2^2} \frac{1}{F_{\alpha/2}(n_1-1, n_2-1)} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{S_1^2}{S_2^2} \frac{1}{F_{1-\alpha/2}(n_1-1, n_2-1)}\right\} = 1 - \alpha$$

于是得 σ_1^2/σ_2^2 的一个置信度为 $1-\alpha$ 的置信区间为

$$\left(\frac{S_1^2}{S_2^2} \frac{1}{F_{\alpha/2}(n_1-1, n_2-1)}, \frac{S_1^2}{S_2^2} \frac{1}{F_{1-\alpha/2}(n_1-1, n_2-1)}\right)$$

区间估计常见情形 非正态总体

1. 使用中心极限定理（大样本）

例 7.27 设总体 X 服从 $(0-1)$ 分布 $B(1, p)$, $0 < p < 1$ 是未知参数, $(X_1, X_2, \dots, X_n)$ 是来自该总体的一个样本, 试求 p 的置信度为 $1-\alpha$ 的置信区间.

解 已知 $E(X)=p, D(X)=p(1-p)$, 记

$$U = \frac{\sum_{i=1}^n X_i - np}{\sqrt{np(1-p)}} = \frac{\sqrt{n}(\bar{X} - p)}{\sqrt{p(1-p)}}$$

则 U 的分布仍与 p 有关, 所以, 直接按上述两例的方法来处理是行不通的. 但由中心极限定理, 当 n 充分大时, U 近似地服从 $N(0, 1)$ 分布, 因此, 当 n 充分大时近似地有

$$P\{|U| < u_{\alpha/2}\} \approx 1-\alpha$$

即

$$P\left\{-u_{\alpha/2} < \frac{\sqrt{n}(\bar{X}-p)}{\sqrt{p(1-p)}} < u_{\alpha/2}\right\} \approx 1-\alpha$$

而等式

$$-u_{\alpha/2} < \frac{\sqrt{n}(\bar{X}-p)}{\sqrt{p(1-p)}} < u_{\alpha/2}$$

等价于不等式

$$ap^2 - bp + c < 0$$

此处 $a = n + u_{\alpha/2}^2, b = 2n\bar{X} + u_{\alpha/2}^2, c = n\bar{X}^2$, 解此不等式得

$$\hat{p}_1 < p < \hat{p}_2$$

回顾：中心极限定理（课本p113）

设随机变量 $X_1, X_2, \dots, X_n, \dots$ 相互独立, 服从同一分布, 且 $E(X_k) = \mu, D(X_k) = \sigma^2 \neq 0 (k = 1, 2, \dots)$,

当 n 很大时, 随机变量 $Y_n = \frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n}\sigma}$ 近似地服从标准正态分布 $N(0, 1)$ 。

区间估计常见情形 非正态总体

2. 待估参数和统计量服从分布已知（可用函数写出来） 课本p162 例7.25

区间估计题目

✓ 16. 设总体 X 服从泊松分布. 参数 $\lambda > 0$ 未知, $(X_1, X_2, \dots, X_n)$ 是来自总体 X 的样本, 试在 n 充分大的条件下, 求参数 λ 的置信度近似为 $1-\alpha$ 的置信区间.

大样本区间估计: λ 同时是均值和方差

23. 随机地从甲批导线中抽取 4 根, 从乙批导线中抽取 5 根, 测得其电阻 (单位: Ω) 为:

甲批导线: 0.143 0.142 0.143 0.137

乙批导线: 0.140 0.142 0.136 0.138 0.140

方差相同, 求 S_w 和 t 分布

设测试数据分别服从正态分布 $N(\mu_1, \sigma^2)$ 和 $N(\mu_2, \sigma^2)$, 试求 $\mu_1 - \mu_2$ 的置信度为 0.95 的置信区间.

24. 从某地区随机地选取男、女各 100 名, 以估计男、女平均身高之差, 测量并计算得男子身高的样本均值为 1.71 m, 样本标准差为 0.035 m, 女子身高的样本均值为 1.67 m, 样本标准差为 0.038 m, 试求男、女身高平均值之差的置信度为 0.95 的置信区间.

大样本: 方差直接加权平均

25. 设 $(X_1, X_2, \dots, X_{n_1})$ 和 $(Y_1, Y_2, \dots, Y_{n_2})$ 是分别来自正态总体 $N(\mu_1, \sigma_1^2)$ 和 $N(\mu_2, \sigma_2^2)$ 的独立样本, 在 μ_1 和 μ_2 都已知的情形, 求方差比 σ_1^2/σ_2^2 的置信度为 $1-\alpha$ 的置信区间、置信下限和置信上限.

F 分布: 用 $(X_i - \mu)^2$ 表示 χ^2

矩估计

已知:

样本矩 A_k (样本统计量) 是总体矩 α_k (总体分布的统计量) 的无偏估计

$(X_1, X_2, \dots, X_n)$ 服从某分布
 PDF: $f_X(x, \theta, \alpha)$
 或者直接由该分布已知的均值方差来表示
 $g_1(\theta, \alpha) = E(X) = \int_{-\infty}^{+\infty} x f_X(x, \theta, \alpha) dx$
 $g_2(\theta, \alpha) = E(X^2) = \int_{-\infty}^{+\infty} x^2 f_X(x, \theta, \alpha) dx$

算矩
 $\alpha_1 = E(X)$
 $\alpha_2 = E(X^2) = D(X) + [E(X)]^2$

矩估计法
 $\hat{\alpha}_1 = A_1$
 $\hat{\alpha}_2 = A_2 = B_2 + A_1^2$
 B_2 : 样本二阶中心矩
 $B_2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^2$

解方程
 $\begin{cases} g_1(\theta, \alpha) = A_1 \\ g_2(\theta, \alpha) = A_2 \end{cases}$
 得 $\hat{\theta} = \dots$
 $\hat{\alpha} = \dots$

例3 设总体 X 的均值 $E(X) = \mu$, 方差 $D(X) = \sigma^2$ 都存在, 且 $\sigma^2 > 0$. 但 μ, σ^2 均为未知. $X_1, X_2, \dots, X_n$ 为来自总体 X 的样本, 求 μ, σ^2 的矩估计量

解:

$$\begin{cases} \alpha_1 = E(X) = \mu \\ \alpha_2 = E(X^2) = D(X) + [E(X)]^2 = \sigma^2 + \mu^2 \end{cases}$$

解得 $\mu = \alpha_1, \sigma^2 = \alpha_2 - \alpha_1^2$

由矩估计法, $\begin{cases} \hat{\mu} = A_1 = \bar{X} \\ \hat{\sigma}^2 = A_2 - A_1^2 \end{cases} \rightarrow \begin{cases} \hat{\mu} = \bar{X}, \\ \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \end{cases}$

对任何总体, 总体均值与方差的矩估计量都不变

最大似然估计

已知 总体 X 概率密度为 $f(x, \theta)$

代入所有样本值 $x_1, x_2, \dots, x_n$ 相乘

$$L(x_1, x_2, \dots, x_n; \theta) = \prod_{i=1}^n f(x_i, \theta)$$

目标: 求 L 极大值 (自变量为 θ)

$$\ln L = \sum_{i=1}^n \ln f(x_i, \theta)$$

多个 θ . (✓)

$$\frac{\partial \ln L}{\partial \theta} = 0$$

解 θ 即可

离散: $P\{X=x_i\} = p_i(\theta)$

似然函数不含样本值的情况: 定义求解, 如何选择参数才能使 L 最大

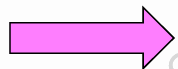
例9 设 $X \sim U(a, b)$, $x_1, x_2, \dots, x_n$ 是 X 的一个样本值, 求 a, b 的极大似然估计值与极大似然估计量.

例8 设总体 $X \sim N(\mu, \sigma^2)$, μ, σ^2 均未知, 又设 $X_1, X_2, \dots, X_n$ 为总体 X 的样本, $x_1, x_2, \dots, x_n$ 为 X 的一组样本观测值, 试求 μ, σ^2 的最大似然估计值及估计量

假设检验

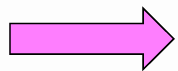
P_s 似乎只要求参数检验?

对参数一无所知



用**参数估计**
的方法处理

若对参数有所了解，但有怀疑猜测需要证实之时



用**假设检验**
的方法处理

参数+样本统计量 服从一定分布 (U , t , F ……)

代入样本观测值

得到一个分布下的值

判断该值是否落入分布的小概率部分内
(α 以上, $1-\alpha$ 以下)
或者说, 偏小/偏大的趋势过于明显

是: 发生了小概率事件, 有理由否定原本的参数
否: 有理由肯定原本参数是对的

假设检验

原假设： H_0 待检验参数 $> = <$ 某个确定值 $\mu < \mu_0$

备择假设： H_1 （在双侧假设检验中，取 H_0 的补集；
在单侧假设检验中，根据实际应用选择）

$$\mu \geq \mu_0$$

显著性水平 α ： **$P\{\text{拒绝}H_0|H_0\text{为真}\}=\alpha$**

假设检验的常见情形 正态总体

1. 单个总体 $N(\mu, \sigma^2)$ 均值 μ 的检验

设总体 $X \sim N(\mu, \sigma^2)$, μ, σ^2 未知, $X_1, X_2, \dots, X_n$ 是来自 X 的样本, 给定显著性水平 α (显著水平)

$$H_0: \mu = \mu_0, \quad H_1: \mu \neq \mu_0$$

1). σ^2 已知 (为 σ_0^2)

取统计量:

$$U = \frac{\bar{X} - \mu_0}{\sigma_0 / \sqrt{n}}$$

—— u 检验法

拒绝域:

$$|u| \geq u_{\alpha/2}$$

2). σ^2 未知

取统计量

$$t = \frac{\bar{X} - \mu_0}{S / \sqrt{n}}$$

—— t 检验法

拒绝域

$$|t| \geq t_{\alpha/2}(n-1)$$

3). σ^2 的检验问题

$$H_0: \sigma^2 = \sigma_0^2, H_1: \sigma^2 \neq \sigma_0^2$$

取检验统计量: $\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$ —— χ^2 检验法

拒绝域: $\chi^2 \leq \chi_{1-\alpha/2}^2(n-1) \quad \chi^2 \geq \chi_{\alpha/2}^2(n-1)$

假设检验的常见情形

正态总体

单边假设检验

2. 两个总体 $N(\mu_1, \sigma_1^2)$ 与 $N(\mu_2, \sigma_2^2)$ 的情形

1) 两总体均值差的检验($\sigma_1^2 = \sigma_2^2$ 未知)

取检验统计量
$$t = \frac{(\bar{X} - \bar{Y}) - \delta}{S_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad \text{—— 检验法}$$

检验假设 $H_0: \mu_1 - \mu_2 = \delta, \quad H_1: \mu_1 - \mu_2 \neq \delta$

拒绝域: $|t| \geq t_{\alpha/2}(n_1 + n_2 - 2)$

课本p190: 拒绝域的设置只与 H_1 有关, 因此单边假设检验的拒绝域和双边的大体相同

$$\alpha/2 \rightarrow \alpha$$

3. 双正态总体方差的检验--F检验

检验 $H_0: \sigma_1^2 = \sigma_2^2, \quad H_1: \sigma_1^2 \neq \sigma_2^2$

取检验统计量 $F = S_1^2 / S_2^2$

拒绝域: $F \geq F_{n_1-1, n_2-1}(\alpha/2) \quad \text{或} \quad F \leq F_{n_1-1, n_2-1}(1-\alpha/2)$

假设检验的常见情形

非正态总体

1. 大样本：使用中心极限定理（课本p182 例8.10） $|u| > u_{\alpha/2}$
2. 小样本的例子：应用结论 $2n\lambda\bar{X} \sim \chi^2(2n)$

成对数据（正态）

1. 可以使用两个总体的方法
2. 也可构造 $Z = Y - X$ ，再使用单个正态总体的t检验法（p195）

假设检验的题目

H_0 , H_1 的选择?

例 8.5 某工厂厂方断言该厂生产的小型电机在正常负载条件下平均消耗电流不会超过 0.8A。现随机抽取 16 台马达,发现它们消耗电流平均是 0.92A,而由这 16 个样本算出的样本标准差为 0.32A,假定这种电机的电流消耗 X 服从正态分布,并且检验水平 $\alpha = 0.05$ 。问根据这一抽样结果,能否否定厂方的断言?

从这个例子,我们看到,随着问题提出的不同(把哪一个断言作为原假设的不同),统计检验的结果得出截然相反的结论。这一点可能使得一些对统计思想不了解的人感到迷惑不解。实际上,这里有个着眼点不同的问题。当把“厂方断言正确”作为原假设时,我们是根据该厂以往的表现和信誉,对其断言已有较大的信任,只有很不利于它的观察结果,才能改变我们的看法,因而难于拒绝这个断言。反之,当把“厂方断言不正确”作为原假设时,我们一开始就对该厂产品抱怀疑态度,只有很有利于该厂的观察结果才能改变我们的看法。因此,在所得观察数据并非决定性地偏于一方时,我们的着眼点决定了所下的结论。

随机过程

离散序列 $\{X(n, \theta), n = 0, 1, 2, \dots\}$

连续时间: $X(t, \omega)$

$$\textcircled{1} F_t(x) = P\{X(t) \leq x\} = P\{\omega \mid X(t, \omega) \leq x\}$$

↳ 固定 t , 在这个时间点上, 满足 $X(t, \omega) \leq x$ 的 ω 的概率.

$$\textcircled{2} m_X(t) = E[X(t)] = \sum_{k=1}^{|\Omega|} p_k \cdot X(t, \omega_k) = \int_{-\infty}^{+\infty} X(t, \omega) f_\omega(\omega) d\omega$$

(离散) (连续)

↳ 固定 t , X 由 ω 控制, $\sum_{\omega \in \Omega} \text{每个 } \omega \text{ 的概率} \cdot \omega \text{ 控制的 } X \text{ 的值}$

$$\textcircled{3} D(t) = D[X(t)] = E[(X(t) - m(t))^2]$$

目录:

1 2 3 概念

4 二维分布

5 泊松过程

6 独立增量过程

7

示出来.

⑨ 设随机过程 $\{X(t), -\infty < t < +\infty\}$ 只有四条样本曲线 $X(t, \omega_1) = 1, X(t, \omega_2) = -1, X(t, \omega_3) = \sin t, X(t, \omega_4) = \cos t$, 且 $P(\{\omega_i\}) = \frac{1}{4}, i = 1, 2, 3, 4$. 求此随机过程的均值函数与自相关函数.

④ 二维的:

(1) $X(t_1), X(t_2)$

分布函数:
$$F_{t_1, t_2}(x_1, x_2) = P\{X(t_1) \leq x_1, X(t_2) \leq x_2\}$$
$$= P\{\omega \mid X(t_1, \omega) \leq x_1, X(t_2, \omega) \leq x_2\}$$

降低难度: 有限个 ω (样本曲线).

自相关函数:
$$R_X(t_1, t_2) = E[X(t_1)X(t_2)]$$
$$= \int_{-\infty}^{+\infty} x(t_1, \omega) x(t_2, \omega) f(\omega) d\omega.$$

自协方差函数:
$$\text{Cov}_X(t_1, t_2) = R_X(t_1, t_2) - m(t_1)m(t_2)$$

⑤

(2) $X(t_1), Y(t_2)$

$$R_{XY}(t_1, t_2) = E[X(t_1)Y(t_2)]$$

$$\text{Cov}_{XY}(t_1, t_2) = E[(X(t_1) - m_X(t_1))(Y(t_2) - m_Y(t_2))]$$

⑤. 泊松过程.

(1) $N(0)=0$

(2) 独立增量



$X(t_{k+1}) - X(t_k)$ 与 $X(t_k) - X(t_{k-1})$ 无关(独立)

区分: 正交增量

(3) $P\{N(s+t) - N(s) = k\} = \frac{(\lambda t)^k}{k!} e^{-\lambda t}$

$E\{[X(t_{k+1}) - X(t_k)][X(t_{k-1}) - X(t_{k-2})]\} = 0.$

泊松过程不是正交增量

布朗运动是

经过 t 个单位时间, 发生的

次数 k 服从 λt 的泊松分布.

* 泊松过程中的等待

设 1 分钟卖出报纸的数量服从 λ 的泊松分布.

问卖出 n 份报纸需要的时间 T 的分布?

• Key:

解: 设 $L_n = T_n - T_{n-1}$ 即卖出第 $n-1$ 与 n 份的间隔.

T_n -已卖出的: 与 L_n 无关.

$$\begin{aligned} P(T_n - T_{n-1} = k) &= P\{\text{前 } (L_n - 1) \text{ 分钟都没卖出, 最后一分钟卖出一份}\} \\ &= \left[\frac{\lambda^0}{0!} e^{-\lambda} \right]^{k-1} \cdot \left[\frac{\lambda^1}{1!} e^{-\lambda} \right] \\ &= e^{-\lambda(k-1)} \cdot \lambda e^{-\lambda} = \lambda e^{-\lambda k} \quad (k > 0) \end{aligned}$$

即 L_n 服从指数分布.

也是 $\alpha=1$, λ 的 Γ 分布

$$L = L_1 + L_2 + \dots + L_n \sim \Gamma(n, \lambda) \quad (\text{可加性})$$

$$E(L) = \frac{n}{\lambda} \quad D(L) = \frac{n}{\lambda^2}$$

泊松过程中的 t ：时间单位化。单位时间的期望为 λ ，于是 t 个单位时间的期望为 λt （泊松可加性+直观理解）

定理1 强度为 λ 的泊松过程的时间间隔 $\{T_n\}$ 是相互独立的随机变量，且服从参数为 λ 的同一个指数分布

定理2 设 $\{N(t), t \geq 0\}$ 是计数过程，若时间间隔 $\{T_n\}$ 是相互独立的随机变量，且服从参数为 λ 同一个指数分布，则 $\{N(t), t \geq 0\}$ 是强度为 λ 的泊松过程

定理1和定理2刻画了泊松过程的本质：一个计数过程是泊松过程，当且仅当时间间隔相互独立且服从同一个指数分布

无记忆性，上一次发生的时间、多少和上一次与这一次的间隔无关（独立）（不相关）

⑥ 独立增量过程的协方差函数 (假定 $N(0)=0$). 第四章.

$$\text{Cov}(N(s), N(t)) = \text{Cov}(N(s), N(t) - N(s) + N(s))$$

$s < t$

$$= \text{Cov}(N(s), N(s)) + \text{Cov}(N(s) - N(0), N(t) - N(s))$$

$$= D(N(s))$$

$$\text{Cov}(x_1 + x_2, Y)$$

$$= \text{Cov}(x_1, Y) + \text{Cov}(x_2, Y)$$

↓
独立. 故 $E(XY) = E(X)E(Y)$

$$R_X(\cdot) = m(\cdot)m(\cdot)$$

总结: $\text{Cov}(N(t_1), N(t_2)) = D(N(\min\{t_1, t_2\}))$

插播：随机游走 (简单版)

随机游走

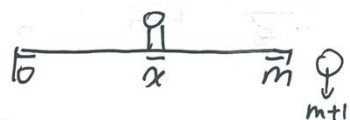
$$\{X_n, n=1, 2, \dots, n\}, \quad P(X_i=1)=p=1-P(X_i=-1)$$

游走者位于位置 x 处，每经过一单位时间，随机向左或向右移 1 个单位。

经过 n 单位时间： $X_n = x + X_1 + X_2 + \dots + X_n$

(1) 到达某个位置 (例如 0) 的概率? $p = \frac{1}{2}$

设 S_n 为初始位置 $x=n$ 时到达边界的概率。



$S(0)=1$ 在任何位置 x_n
 $\frac{1}{2}$ 向左, $\frac{1}{2}$ 向右

$\therefore S_n = \frac{1}{2}S_{n-1} + \frac{1}{2}S_{n+1}$ 递推推式子.

得 $S_n = kn + S_0 = (S_1 - S_0)n + S_0$

补充右边界 $S_{m+1}=0$ 则 $k = \frac{1}{m+1}$ $S_n = \frac{n}{m+1} + 1$

$S_n = 1 - \frac{n}{m+1}$ 也可记作从 $x=n$ 开始, 到 $m+1$ 边界的概率为 $\frac{n}{m+1}$

⑦ 平稳过程

$E[X(t)]$ 与 $E[X^2(t)]$ 都存在。
于是该求的都能求(?)

宽平稳: $X_T = \{X(t), t \in T\}$ 是二阶矩过程

且在任何 t 下, $E[X(t)]$ 是常数(而不是与 t 有关的函数)

任何 τ 下, $R_X(\tau) = E[X(t)X(t+\tau)]$ 与 t 无关。

正态白噪声: 互不相关的序列 $\{X(n)\}$:
 $X(n) = N(0, \sigma^2), n=0, 1, 2, \dots$

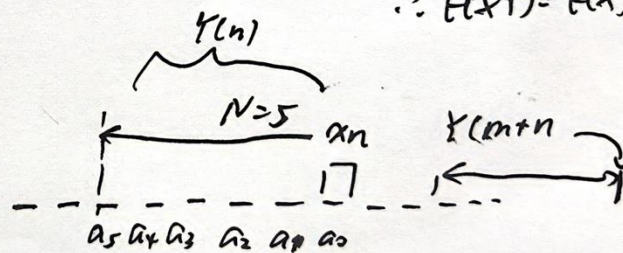
$$E[X(n)] = 0$$

$$R_X(n) = E[X(m)X(m+n)] = \begin{cases} \sigma^2 & n=0 \\ 0 & n \neq 0 \end{cases}$$

同一个: 自己的方差
不相关: 正态不相关 $\Leftrightarrow$ 独立
 $\therefore E(XY) = E(X)E(Y)$

滑动和

$$Y(n) = \sum_{k=0}^N a_k X(n-k)$$



$$R_Y(m) = 0 \quad |m| > N \quad \text{滑动长度超过 } Y \text{ 的跨度}$$

宽平稳过程的条件旨在说明

1. 过程的均值不随时间而改变
2. 过程前后的相关性只与间隔的距离有关

补: 随机电报信号过程

8 自相关函数:

“平稳”: $R_X(\tau) = E[X(t)X(t+\tau)]$

① $R_X(\tau) \geq 0$

② $|R_X(\tau)| \leq R_X(0)$ 自己与自己最相关. $a \cdot b \leq \sqrt{a \cdot a} \sqrt{b \cdot b}$

证明: 柯西-施瓦茨不等式 $[\int f(x)g(x)]^2 \leq \int f^2(x) \int g^2(x)$

$$[E(XY)]^2 \leq E(X^2) E(Y^2)$$

$$E[X(t)X(t+\tau)] \leq \sqrt{E[X^2(t)]} \sqrt{E[X^2(t+\tau)]}$$

$$= R_X(0)$$

③ 偶函数 (易见)

④ 正定 略. 不会证.

9 互相关函数.

$X(t), Y(t)$

平稳相关: $R_{XY}(\tau) = E[X(t)Y(t+\tau)]$ 与 t 无关.

$$C_{XY}(\tau) = R_{XY}(\tau) - m_X m_Y$$

① $R_{XY}(-\tau) = E[X(t)Y(t-\tau)]$

$$= E[Y(t-\tau)X[(t-\tau)+\tau]] = R_{YX}(\tau)$$

② $|R_{XY}(\tau)| \leq \sqrt{R_X(0)} \sqrt{R_Y(0)}$

练习题

9. 设 A, B, Θ 是相互独立的随机变量, $E(A) = a, E(B) = b, D(A) = \sigma_1^2, D(B) = \sigma_2^2, \Theta \sim U[-\pi, \pi], \omega_0$ 为常数, $X(t) = A \sin(\omega_0 t + \Theta), Y(t) = B \cos(\omega_0 t + \Theta)$, 试证明 $\{X(t), -\infty < t < +\infty\}$ 与 $\{Y(t), -\infty < t < +\infty\}$ 是平稳相关的平稳过程.

傅里叶变换的理解

傅里叶变换

高数里学的: $f(x) = \frac{a_0}{2} + \sum_{k=1}^{+\infty} (a_k \cos kx + b_k \sin kx)$

高数里怎么转换: $\cos kx = \frac{e^{ikx} + e^{-ikx}}{2}$, $\sin kx = \frac{e^{ikx} - e^{-ikx}}{2i} = \frac{(e^{ikx} - e^{-ikx})i}{2}$ ($i^2 = -1$)

于是 $f(x) = C_0 + \sum_{k=1}^{+\infty} \frac{e^{ikx}}{2} (a_k - ib_k) + \frac{e^{-ikx}}{2} (a_k + ib_k)$
 $= C_0 + \sum_{k=1}^{+\infty} (C_{+k} e^{ikx} + C_{-k} e^{-ikx})$

其中 $C_{\pm k} = \frac{a_k \mp ib_k}{2}$

概心: (连续)

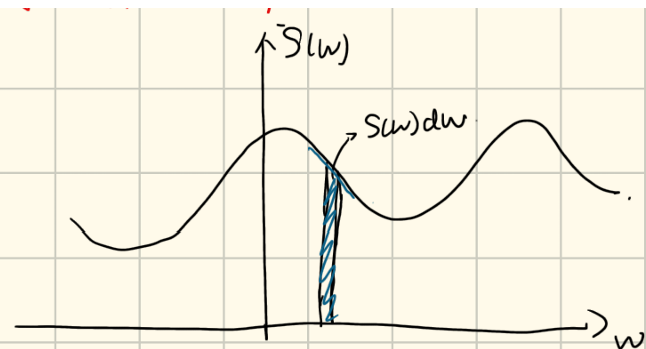
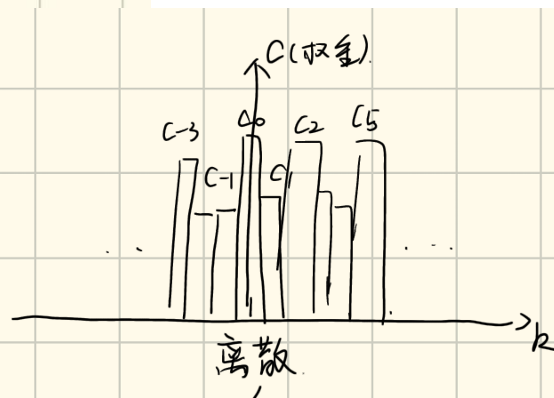
一个信号 $x(t)$

$x(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} S(\omega) e^{i\omega t} d\omega$ (反变换)

$x(t) \leftrightarrow f(x)$: 被分解的函数 ($R_x(t)$)

$\omega \leftrightarrow k$: 频率

本质: 将一个信号 $f(x)$ 拆成无数个以 $\cos(kx), \sin(kx)$ 为基的线性加权和。在谱密度里, k 之间的间隔是无穷小, 因此用“谱密度”来表示 (类比概率密度和概率分布函数)



(P282) $\hat{F}(\omega) = \int S(\omega) d\omega + C$

类比: 概率密度函数 $f(x)$
 与 $F(x)$ 分布函数

傅里叶变换的理解

线性代数: 向量内积, 投影

$$a \cdot b = \|a\| \|b\| \cos \langle a, b \rangle$$

如果 e_1, e_2, e_3 是一组 $\mathbb{R}^3$ 内的基向量

则 a 在该组基下的坐标可求

$$\vec{a} = [\vec{e}_1, \vec{e}_2, \vec{e}_3] \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$$
$$= a_1 \vec{e}_1 + a_2 \vec{e}_2 + a_3 \vec{e}_3$$

正交性:

$$a \cdot \vec{e}_1 = a_1 \vec{e}_1 \cdot \vec{e}_1$$

$$a_1 = \frac{a \cdot \vec{e}_1}{\|\vec{e}_1\|^2}$$

傅里叶变换的实质是求函数在一组指定函数正交基下的坐标 (基的个数有无穷个)

函数的内积

$f, g \in C[a, b]$ 在这个函数空间内 (叫什么什么空间? 忘了)

$$\langle f, g \rangle = \int_a^b f(x) g(x) dx$$

设有一组正交基 $g_1, g_2, \dots, g_n$

$$\int_a^b g_i g_k = 0 \quad \int_a^b g_i g_i = 2\pi$$

$$\downarrow$$
$$\langle g_i, g_k \rangle = 0$$

$$\downarrow$$
$$\|g_i\|^2 = 2\pi$$

则 f 在每个 g 上的投影坐标

$$a_i = \frac{\langle f, g_i \rangle}{\|g_i\|^2} = \frac{1}{2\pi} \int_a^b f(x) g_i(x) dx$$

让基的个数趋于无穷, $[a, b]$ 变成 $(-\infty, +\infty)$

则得到傅里叶变换中求

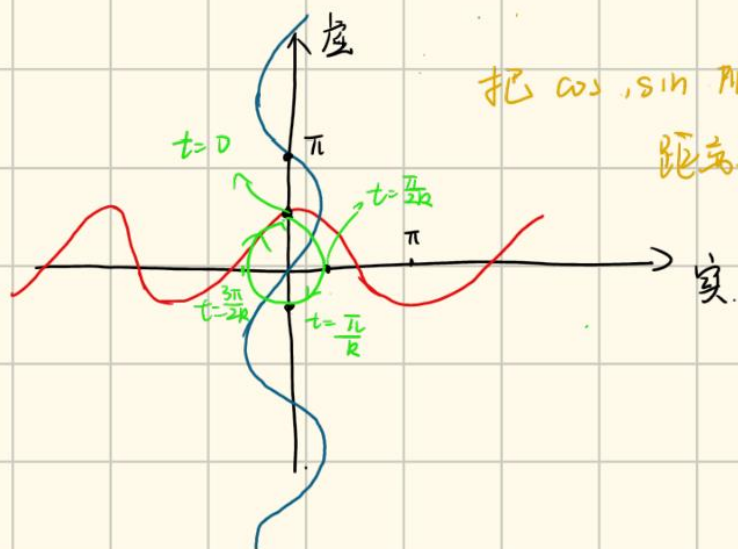
a_k, b_k, C_k 的方法

傅里叶变换的理解

取正交基为 $e^{-ik\tau}$ $k=0, \pm 1, \pm 2, \dots$

(总之 $e^{\pm ik\tau}$ 就是正交的, 不会证, 记住)

$$e^{-ik\tau} = \cos k\tau - i \sin k\tau$$



把 $\cos, \sin$ 顺着 t 向前进的时刻

距离各自的轴的 距离 " 度 " 到

一个圆上

求功率谱密度/谱密度/ $S(w)$ /自谱密度

$$\text{则 } S(k) = \int_{-\infty}^{+\infty} R_x(\tau) e^{-ik\tau} d\tau$$

$k \rightarrow$ 连续的 w . 得到 课本上的正变换.

$$S(w) = \int_{-\infty}^{+\infty} R_x(\tau) e^{-i w \tau} d\tau \stackrel{\text{def}}{=} \mathcal{F}[R_x(\tau)]$$

$$R_x(\tau) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} S_x(w) e^{i w \tau} dw \stackrel{\text{def}}{=} \mathcal{F}^{-1}[S_x(w)]$$

⑩ 平稳过程 $x(t)$ 的平均功率:

$$\lim_{T \rightarrow +\infty} E \left\{ \frac{1}{2T} \int_{-T}^T x^2(t) dt \right\} = \frac{1}{2\pi} \int_{-\infty}^{+\infty} S(\omega) d\omega$$
$$= R_x(0) \quad \text{均方值}$$

平均功率 (物理) 这一部分似乎不是重点?

⑪ $S(\omega)$ 性质:

$S_x(\omega)$ 是 ω 的实的偶函数且非负.

课本简单象征性看一下, 然后跳过,
(然后发现自己什么都没记住)

互谱密度:

$$(1) S_{xy}(\omega) = \int_{-\infty}^{+\infty} R_{xy}(\tau) e^{-i\omega\tau} d\tau$$

$\rightarrow x, y$ 的互相关函数 (要求两者平稳相关)
且 x, y 都是平稳过程.

$$(2) \text{共轭对称: } S_{xy}(-\omega) = S_{xy}^*(\omega) = S_{yx}(\omega)$$

$$\cancel{S_{xy}(\omega) = S_{yx}(\omega)} \quad (\text{为什么不起呢?})$$

$$(3) \text{柯西-施瓦茨: } |S_{xy}(\omega)| \leq \sqrt{S_x(\omega)} \sqrt{S_y(\omega)}$$

(4) 互谱密度不一定是实的, 具有实部 (偶函数) 和虚部 (奇函数)

怎么不查表求 $S(w)$

以表为例（如果真的出了特别难积的那我只能原地爆炸）

① 最普通 $e^{-\alpha|\tau|}$

$$R_x(\tau) = \sigma^2 e^{-\alpha|\tau|}$$

路线1: 合并e

$$S_x(w) = \int_{-\infty}^{+\infty} \sigma^2 e^{-\alpha|\tau|} e^{-i w \tau} d\tau$$

$$= \sigma^2 \int_0^{+\infty} e^{-(\alpha + i w) \tau} d\tau + \sigma^2 \int_{-\infty}^0 e^{(\alpha - i w) \tau} d\tau$$

$$= \frac{\sigma^2}{-(\alpha + i w)} e^{-(\alpha + i w) \tau} \Big|_0^{+\infty} + \frac{\sigma^2}{\alpha - i w} e^{(\alpha - i w) \tau} \Big|_{-\infty}^0$$

$$= \frac{\sigma^2}{\alpha + i w} + \frac{\sigma^2}{\alpha - i w} = \frac{2\alpha \sigma^2}{\alpha^2 + w^2}$$

怎么不查表求 $S(w)$

以表为例（如果真的出了特别难积的那我只能原地爆炸）

② 低通白噪声

$$S_{xx}(w) = \begin{cases} G^2 & |w| \leq w_0 \\ 0 & |w| > w_0 \end{cases}$$

频率小于 w_0 时通过

$$R_{xx}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} G^2 e^{i w \tau} dw$$

$$= \frac{G^2}{2\pi} \int_{-w_0}^{w_0} (\cos w \tau + i \sin w \tau) dw$$

↑奇

(直接积分不出来)

$$= \frac{G^2}{2\pi} \tau w_0^2 \sin w_0 \tau$$

$$= \frac{G^2 \sin w_0 \tau}{\pi \tau} = \frac{G^2 w_0}{\pi} \left(\frac{\sin w_0 \tau}{w_0 \tau} \right) \quad \tau \neq 0$$

$$\tau = 0 = \frac{G^2 w_0}{\pi}$$

路线2: $e^{i w \tau}$ 拆成三角函数积分

怎么不查表求 $S(w)$

路线3：把三角函数
拆成 $e^{+/-i\omega t}$

$$\begin{aligned} R_x(\tau) &= b^2 e^{-\alpha|\tau|} \cos(\omega_0 \tau) \\ S_x(\omega) &= b^2 \int_{-\infty}^{+\infty} e^{-\alpha|\tau|} \frac{e^{i\omega_0 \tau} + e^{-i\omega_0 \tau}}{2} e^{-i\omega \tau} d\tau \quad \text{再合并.} \\ &= \frac{b^2}{2} \int_{-\infty}^{+\infty} e^{-\alpha|\tau|} \left[e^{-i(\omega - \omega_0)\tau} + e^{-i(\omega + \omega_0)\tau} \right] d\tau \\ &= \frac{b^2}{2} \left[\frac{2\alpha}{(\omega - \omega_0)^2 + \alpha^2} + \frac{2\alpha}{(\omega + \omega_0)^2 + \alpha^2} \right] \\ &= \alpha b^2 \left[\frac{1}{(\omega - \omega_0)^2 + \alpha^2} + \frac{1}{(\omega + \omega_0)^2 + \alpha^2} \right] \end{aligned}$$

怎么不查表求 $S(w)$

⑥ δ 狄拉克.

$$\int_{-\infty}^{+\infty} \delta(\tau) e^{-i\omega\tau} d\tau = 1$$

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} \delta(\tau) e^{i\omega\tau} d\tau = \frac{1}{2\pi}$$

狄拉克函数的特点：书p289

似乎只需要知道筛选性就可以了

若 $S_x(\omega) = 1$ (常数).

则 $R_x(\tau) = \delta(\tau)$

若 $R_x(\tau) = \delta^2$ (常数).

则 $S_x(\omega) = 2\pi \delta(\omega)$ (只在 $\omega=0$ 有一个脉冲)

为什么引入狄拉克:

p 290

$$= \int_{-\infty}^{+\infty} \delta^2 \cos(\omega_0\tau) e^{-i\omega\tau} d\tau$$

$$= \delta^2 \int_{-\infty}^{+\infty} (e^{i(\omega_0-\omega)\tau} + e^{-i(\omega_0+\omega)\tau}) d\tau \quad \text{不可积.}$$

怎么不查表求S(w)

练习题

$$R_x(\tau) = \begin{cases} 1 - \frac{|\tau|}{T} & |\tau| \leq T \\ 0 & |\tau| > T \end{cases}$$

各态历经性

1. 直接判断 (最推荐)

定义1 设 $\{X(t), -\infty < t < +\infty\}$ 为随机过程

$$\langle X(t) \rangle \triangleq (\text{m.s.}) \lim_{T \rightarrow +\infty} \frac{1}{2T} \int_{-T}^T X(t) dt$$

$$\langle X(t+\tau) X(t) \rangle \triangleq (\text{m.s.}) \lim_{T \rightarrow +\infty} \frac{1}{2T} \int_{-T}^T X(t+\tau) X(t) dt$$

如果上面两个 $\langle \rangle$ 分别等于
 $m(t), R(t)$: 是各态历经过程

例: $X(t) = a \cos(\omega_0 t + \phi)$, a, ω_0 是常数. $\theta \sim U(0, 2\pi)$.

$$\text{时间均值: } \langle X(t) \rangle = \lim_{T \rightarrow +\infty} \frac{1}{2T} \int_{-T}^T X(t) dt$$

$$= \lim_{T \rightarrow +\infty} \frac{1}{2T} \int_{-T}^T a \cos(\omega_0 t + \phi) dt$$

$$= \lim_{T \rightarrow +\infty} \frac{a \sin(\omega_0 T) \cos \phi}{\omega_0 T} = 0$$

$$\text{时间相关函数: } \langle X(t) X(t+\tau) \rangle = \lim_{T \rightarrow +\infty} \frac{1}{2T} \int_{-T}^T X(t) X(t+\tau) dt$$

$$= \lim_{T \rightarrow +\infty} \frac{1}{2T} \int_{-T}^T a^2 \cos(\omega_0 t + \phi) \cos(\omega_0 t + \omega_0 \tau + \phi) dt$$

$$= \lim_{T \rightarrow +\infty} \frac{a^2}{2T} \int_{-T}^T \frac{1}{2} (\cos(2\omega_0 t + \omega_0 \tau + \phi) + \cos \omega_0 \tau) dt$$

$$= \frac{a^2}{2} \cos \omega_0 \tau \quad (\text{第一次被分母 } T \rightarrow +\infty \text{ 消掉})$$

$$\text{均值 } m_X(t) = 0$$

$$R_X(\tau) = \frac{a^2}{2} \cos \omega_0 \tau$$

$\therefore$ 是平稳过程.

且有各态历经性.

各态历经性

2. 使用定理（不推荐）

定理1 数学期望的遍历性定理

设 $\{X(t), -\infty < t < +\infty\}$ 为均方连续平稳过程，则 $X(t)$ 有均值**各态历经性**的充要条件为

$$\lim_{T \rightarrow +\infty} \frac{1}{2T} \int_{-2T}^{2T} \left(1 - \frac{|\tau|}{2T}\right) (B(\tau) - m^2) d\tau = 0$$

其中 $E[X(t)] = m$, $B(\tau)$ 是**相关函数**, $\tau = t_1 - t_2$

定理2 自相关函数**各态历经性**定理

设 $\{X(t), -\infty < t < +\infty\}$ 是均方连续平稳过程 $Z = X(t + \lambda) X(t)$

则相关函数 $B_Z(\tau)$ 具有**各态历经性**的充要条件为

$$\lim_{T \rightarrow +\infty} \frac{1}{2T} \int_{-2T}^{2T} \left(1 - \frac{|\tau|}{2T}\right) (B_Z(\tau) - m_Z^2) d\tau = 0$$

其中 $B_Z(\tau) = E[X(t + \lambda + \tau) X(t + \tau) X(t + \lambda) X(t)]$

$m_Z = E[X(t + \lambda) X(t)]$

各态历经性

3. 使用定理的变形 (推荐)

$$\lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^{2T} \left(1 - \frac{\tau}{2T}\right) [R_X(\tau) - m_X^2] d\tau = 0$$

$$\text{即 } \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^{2T} \left(1 - \frac{\tau}{2T}\right) C_X(\tau) d\tau = 0$$

方法 3: 判断 $C_X(\tau) = 0$. (平均值)

理解: 当 τ 无限大时; $X(t)$ 与 $X(t+\tau)$ 不线性相关.

4. (自相关函数).

$$\lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^{2T} \left(1 - \frac{\tau_1}{2T}\right) [R_T(\tau_1) - R_X^2(\tau)] d\tau_1 = 0$$

$$Y = X(t)X(t+\tau) \quad E(Y) = R_X(\tau)$$

$$\begin{aligned} R_{YY}(\tau) &= E[X(t)X(t+\tau)X(t+\tau_1)X(t+\tau+\tau_1)] \\ &= R_X(\tau_1) \cdot (\text{定义}) \end{aligned}$$

即判断 Y 的 均值函数 是否具有各态历经性 (转到方法 3)

Thanks!