概率论与数理统计笔记

衔瑜[1]

2021 年 12 月 17 日

[1]Email: fish233yeah@163.com

目录

Chapter 1

知识点总结

1.1 随机事件及其概率

1.1.1 基本知识

1. **样本空间与样本点**：对于随机事件E，它的所有可能结果组成的集合即为E的样本空间，通常记为S；而样本空间中的元素（即E的每个结果）则称为样本点

2. **摩根率（对偶率）**：$\overline{A+B}=\overline{A}\,\overline{B}$，$\overline{AB}=\overline{A}+\overline{B}$

3. **排列组合公式**：$A_n^m=\frac{n!}{(n-m)!}$，$C_n^m=\frac{A_n^m}{m!}$。此外我们有：$C_n^m=C_n^{n-m}$和$C_n^m=C_{n-1}^m+C_{n-1}^{m-1}$

4. **条件概率**：称$P(B|A)=\frac{P(AB)}{P(A)}$为事件$B$在事件$A$发生的条件下发生的概率。而对于多个事件$B_1,B_2,\cdots$，若它们为两两互不相容的事件，则我们有$P\left(\bigcup\limits_{i=1}^{\infty}B_i|A\right)=\sum\limits_{i=1}^{\infty}P(B_i|A)$

5. **全概率公式**：设E的样本空间为S，A为E的事件，$B_1,B_2,\cdots,B_n$为S的一个划分，且$P(B_i)>0$，则有：
$$P(A)=P(A|B_1)P(B_1)+P(A|B_2)P(B_2)+\cdots+P(A|B_n)P(B_n)$$

6. **贝叶斯公式**：设E的样本空间为S，A为E的事件，$B_1,B_2,\cdots,B_n$为S的一个划分，且$P(A)>0$，$P(B_i)>0$，则有：
$$P(B_i|A)=\frac{P(A|B_i)P(B_i)}{\sum\limits_{j=1}^{n}P\left(A|B_j\right)P\left(B_j\right)}$$

7. **事件独立定义**：设$A_1,A_2,\cdots,A_n$为n个事件，若对于其中任意2个、任意3个、$\cdots$、任意n个事件的积事件的概率都等于各事件概率之积，则称事件$A_1,A_2,\cdots,A_n$相互独立

8. 若非负实数$k_1,k_2,\cdots,k_n$满足$k_1+k_2+\cdots+k_n=1$，则分布函数$F_1(x),F_2(x),\cdots,F_n(x)$的线性组合$F(x)=k_1F_1(x)+k_2F_2(x)+\cdots+k_nF_n(x)$仍为某一随机变量的分布函数

1.2 计算事件的概率

1.2.1 部分公式

$P(A-B)=P(A)-P(AB)$

$P(A+B)=P(A)+P(B)-P(AB)$

$P(A+B+C)=P(A)+P(B)+P(C)-P(AB)-P(AC)-P(BC)+P(ABC)$

$P(A_1+A_2+\cdots+A_n)=\sum_{i=1}^{n}P(A_i)-\sum_{1\le i<j\le n}P\left(A_iA_j\right)+\cdots+(-1)^{n-1}P(A_1A_2\cdots A_n)$

$P(AB)=P(A)P(B|A)=P(B)P(A|B)$

$P(A_1A_2\cdots A_n)=P(A_1)P(A_2|A_1)P(A_3|A_1A_2)\cdots P(A_n|A_1A_2\cdots A_{n-1})$

1.3 随机变量及其分布

1.3.1 几种离散型随机变量分布律

1. **两点分布**：又称为伯努利分布、$(0-1)$分布，随机变量X只可能取0和1两个值，分布律为：

$$P\{X=k\}=p^k(1-p)^{1-k},\quad k=0,1$$

2. **二项分布**：设试验E只有两个可能的结果A和$\overline{A}$，$P(A)=p$，则称E为伯努利试验，我们将试验E独立重复n次的话则称其为n重伯努利试验。现在假设我们用X表示n重伯努利试验中事件A发生的次数，则其符合二项分布，记为$X\sim b(n,p)$或$X\sim B(n,p)$，分布律为：

$$P\{X=k\}=C_n^kp^k(1-p)^{n-k},\quad k=0,1,2,\cdots,n$$

若有$X\sim B(m,p)$，$Y\sim B(n,p)$，且X和Y相互独立，则$X+Y\sim B(m+n,p)$

3. **泊松分布**：设随机变量X所有可能取的值为$0,1,2,\cdots$，分布率为：

$$P\{X=k\}=\frac{\lambda^ke^{-\lambda}}{k!},\quad k=0,1,2,\cdots$$

其中$\lambda>0$，此时我们称随机变量X服从参数为λ的泊松分布，记为$X\sim\pi(\lambda)$或$X\sim P(\lambda)$。此外，若我们记$np_n=\lambda$，我们还有泊松定理：

$$\lim_{n\to\infty}C_n^kp_n^k(1-p_n)^{n-k}=\frac{\lambda^ke^{-\lambda}}{k!}$$

4. **几何分布**：设在伯努利试验中事件A发生的概率为p，若记事件A首次发生时的经历的试验总次数为X，则我们称随机变量X服从参数为p的几何分布，分布律为：

$$P\{X=k\}=pq^{k-1},\quad k=1,2,\cdots$$

5. **超几何分布**：设在N个物品中有N_1个次品，从中任取$n(1 \le n \le N_1)$个，设X为这n个物品中的次品的数量，则随机变量X服从超几何分布，分布律为：

$$P\{X = k\} = \frac{C_{N_1}^{k} C_{N-N_1}^{n-k}}{C_N^n}$$

1.3.2 几种连续型随机变量分布律

1. **均匀分布**：记为$X \sim U(a,b)$

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & a < x < b \\ 0, & else \end{cases}$$

$$F(x) = \begin{cases} 0, & x < a \\ \dfrac{x-a}{b-a}, & a \le x < b \\ 1, & x \ge b \end{cases}$$

2. **指数分布**：记服从参数$\lambda > 0$的指数分布为$X \sim E(\lambda)$或$X \sim Exp(\lambda)$

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \ge 0 \\ 0, & x < 0 \end{cases}$$

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \ge 0 \\ 0, & x < 0 \end{cases}$$

指数分布具有无记忆性，即：

$$P(X > s + t | X > s) = P(X > t)$$

3. **正态分布**：又叫做高斯分布，记服从参数$\mu, \sigma\,(\sigma > 0)$的正态分布为$X \sim N\left(\mu, \sigma^2\right)$

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

$$F(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt, \quad -\infty < x < \infty$$

当$\mu = 0$，$\sigma = 1$时称随机变量X服从标准正态分布，其概率密度函数和分布函数为

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad -\infty < x < \infty$$

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt, \quad -\infty < x < \infty$$

正态分布有如下几个性质：

(a) 正态分布$N\left(\mu, \sigma^2\right)$的数学期望为$\mu$，方差为$\sigma^2$

(b) 若$X \sim N\left(\mu, \sigma^2\right)$，则$Y = \frac{X-\mu}{\sigma} \sim N(0,1)$

(c) 若$X \sim N\left(\mu, \sigma^2\right)$，则$Y = aX + b \sim N\left(a\mu + b, (a\sigma)^2\right)$, $a \neq 0$

(d) 若$X_i \sim N\left(\mu_i, \sigma_i^2\right)$, $(i = 1, 2, \cdots, n)$，且$X_1, X_2, \cdots, X_n$相互独立，则其线性组合$Y = a_1X_1 \pm a_2X_2 \pm \cdots \pm a_nX_n$仍服从正态分布，且有

$$Y \sim N\left(a_1\mu_1 \pm a_2\mu_2 \pm \cdots \pm a_n\mu_n, a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \cdots + a_n^2\sigma_n^2\right)$$

1.3.3 完整概率分布表

分布	参数	分布律或概率密度函数	期望	方差
(0-1)分布	$0 < p < 1$	$P\{X = k\} = p^k(1-p)^{1-k}$, $k = 0, 1$	p	$p(1-p)$
二项分布	$n \geq 1$ $0 < p < 1$	$P\{X = k\} = C_n^k p^k (1-p)^{n-k}$ $k = 0, 1, 2, \cdots, n$	np	$np(1-p)$
负二项分布 (帕斯卡分布)	$r \geq 1$ $0 < p < 1$	$P\{X = k\} = C_{k-1}^{r-1} p^r (1-p)^{k-r}$ $k = r, r+1, \cdots$	$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$
几何分布	$0 < p < 1$	$P\{X = k\} = p(1-p)^{k-1}$ $k = 1, 2, \cdots$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
超几何分布	$N,\ M,\ n$ $(M \leq N)$ $(n \leq N)$	$P\{X = k\} = \frac{C_M^k \cdot C_{N-M}^{n-k}}{C_N^k}$ $\max\{0, n-N+M\} \leq k \leq \min\{n, M\}$	$\frac{nM}{N}$	$\frac{nM}{N}\left(1 - \frac{M}{N}\right)\left(\frac{N-n}{N-1}\right)$
泊松分布	$\lambda > 0$	$P\{X = k\} = \frac{\lambda^k e^{-\lambda}}{k!}$ $k = 0, 1, 2, \cdots$	λ	λ
均匀分布	$a < b$	$f(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & else \end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
正态分布	μ $\sigma > 0$	$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2
Γ分布	$\alpha > 0$ $\beta > 0$	$f(x) = \begin{cases} \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}}, & x > 0 \\ 0, & else \end{cases}$	$\alpha\beta$	$\alpha\beta^2$
指数分布 (负指数分布)	$\lambda > 0$	$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
χ^2分布	$n \geq 1$	$f(x) = \begin{cases} \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, & x > 0 \\ 0, & else \end{cases}$	n	$2n$
韦布尔分布	$\eta > 0$ $\beta > 0$	$f(x) = \begin{cases} \frac{\beta}{\eta}\left(\frac{x}{\eta}\right)^{\beta-1} e^{-\left(\frac{x}{\eta}\right)^\beta}, & x > 0 \\ 0, & else \end{cases}$	$\eta\Gamma\left(\frac{1}{\beta} + 1\right)$	$\eta^2\left\{\Gamma\left(\frac{2}{\beta} + 1\right) - \left[\Gamma\left(\frac{1}{\beta} + 1\right)\right]^2\right\}$
瑞利分布	$\sigma > 0$	$f(x) = \begin{cases} \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}, & x > 0 \\ 0, & else \end{cases}$	$\sqrt{\frac{\pi}{2}}\sigma$	$\frac{4-\pi}{2}\sigma^2$
β分布	$\alpha > 0$ $\beta > 0$	$f(x) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, & 0 < x < 1 \\ 0, & else \end{cases}$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$

对数正态分布	μ $\sigma>0$	$f(x)=\begin{cases}\frac{1}{\sqrt{2\pi}\sigma x}e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}}, & x>0\\ 0, & else\end{cases}$	$e^{\mu+\frac{\sigma^2}{2}}$	$e^{2\mu+\sigma^2}\left(e^{\sigma^2}-1\right)$
柯西分布	a $\lambda>0$	$f(x)=\frac{1}{\pi}\cdot\frac{\lambda}{\lambda^2+(x-a)^2}$	不存在	不存在
t分布	$n\geq 1$	$f(x)=\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\cdot\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}$	$0,n>1$	$\frac{n}{n-2},n>2$
F分布	n_1,n_2	$f(x)=\begin{cases}\frac{\Gamma\left(\frac{n_1+n_2}{2}\right)\cdot\left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}}\cdot x^{\frac{n_1}{2}-1}}{\Gamma\left(\frac{n_1}{2}\right)\cdot\Gamma\left(\frac{n_2}{2}\right)\cdot\left(1+\frac{n_1}{n_2}x\right)^{\frac{n_1+n_2}{2}}}, & x>0\\ 0, & else\end{cases}$	$\frac{n_2}{n_2-2}$ $n_2>2$	$\frac{2n_2^2(n_1+n_2-2)}{n_1(n_2-2)^2(n_2-4)}$ $n_2>4$

表 1.1: 概率分布表

1.4 多维随机变量及其分布

1.4.1 边缘分布

1. 离散型随机变量边缘分布：

$$p_{i\cdot}=P\{X=x_i\}=\sum_{j=1}^{\infty}p_{ij}$$

$$p_{\cdot j}=P\left\{Y=y_j\right\}=\sum_{i=1}^{\infty}p_{ij}$$

2. 连续型随机变量边缘分布：

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

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

1.4.2 条件分布

1. 离散型随机变量条件分布：

$$P\left\{X=x_i|Y=y_j\right\}=\frac{P\left\{X=x_i,Y=y_j\right\}}{P\left\{Y=y_j\right\}}=\frac{p_{ij}}{p_{\cdot j}}$$

$$P\left\{Y=y_j|X=x_i\right\}=\frac{P\left\{X=x_i,Y=y_j\right\}}{P\{X=x_i\}}=\frac{p_{ij}}{p_{i\cdot}}$$

2. 连续型随机变量条件分布：

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}$$

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}$$

1.4.3 二维随机变量独立性

1. 设$F(x,y)$及$F_X(x)$，$F_Y(y)$分别是二维随机变量(X,Y)的分布函数及边缘分布函数，则随机变量X和Y相互独立的两个相互等价的条件为：

   (a) $F(x,y) = F_X(x) F_Y(y)$在平面上恒成立

   (b) $f(x,y) = f_X(x) f_Y(y)$在平面上几乎处处成立

   当随机变量X和Y相互独立时，显然我们有：

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)} = \frac{f_X(x) f_Y(y)}{f_Y(y)} = f_X(x)$$

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{f_X(x) f_Y(y)}{f_X(x)} = f_Y(y)$$

   即二维随机变量X,Y相互独立的充要条件为X,Y的所有条件分布都与其边缘分布相同

2. 设多维随机变量$(X_1, X_2, \cdots, X_m)$和$(Y_1, Y_2, \cdots, Y_n)$相互独立，则X_i和Y_j相互独立。又若h，g是连续函数，则$h(X_1, X_2, \cdots, X_m)$和$g(Y_1, Y_2, \cdots, Y_n)$相互独立

3. 独立一定不相关，不相关不一定独立

1.4.4 二维随机变量函数$Z = g(X,Y)$的分布

1. 离散型随机变量函数概率分布计算：

$$P(Z = z_k) = P(g(X,Y) = z_k) = \sum_{g\left(x_i, y_j\right) = z_k} p_{ij}$$

2. 连续型随机变量函数概率分布计算：

$$P(Z = z) = P(g(X,Y) \le z) = P((X,Y) \in G) = \iint_G f(x,y)\, dxdy$$

   其中$G = \{(x,y) | g(x,y) \le z\}$

1.4.5 二维均匀分布及二维正态分布的性质

1. **二维均匀分布**

   (a) 定义：
   设G为平面上的有界区域，面积为$S(G)$，若二维随机变量(X,Y)在G上服从均匀分布，则

   $$f(x,y)=\begin{cases}\dfrac{1}{S(G)}, & (x,y)\in G\\ 0, & else\end{cases}$$

   记为$(X,Y)\sim U(G)$

   (b) 性质：

   i. 设矩形区域为$G=\{(x,y)\,|a\le x\le b, c\le y\le d\}$，则二维随机变量$(X,Y)$在矩形区域$G$上服从均匀分布的充分必要条件为$X,Y$分别在$(a,b),(c,d)$上分别服从均匀分布

2. **二维正态分布**

   (a) 定义：
   二维随机变量(X,Y)在平面上服从二维正态分布，则

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

   记为$(X,Y)\sim N\left(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho\right)$

   (b) 性质：

   i. 对于二维正态随机变量(X,Y)，X和Y相互独立的充分必要条件为$\rho=0$

   ii. 二维正态随机变量$(X,Y)\sim N\left(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho\right)$的边缘密度函数$f_X(x)$与$f_Y(y)$与$\rho$的取值无关，恒为

   $$f_X(x)=\frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}$$

   $$f_Y(y)=\frac{1}{\sqrt{2\pi}\sigma_2}e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}$$

   即若(X,Y)服从二维正态分布，则X和Y都服从一维正态分布

   iii. 若X和Y都服从一维正态分布，且X和Y相互独立，则(X,Y)服从二维正态分布

   iv. (X,Y)服从二维正态分布的充分必要条件为X和Y的任意线性组合l_1X+l_2Y（l_1,l_2不全为零）均服从一维正态分布

   v. （正态变量线性变换不变性）若(X,Y)服从二维正态分布，记随机变量

   $$U=aX+bY,\ V=cX+dY,\ 且\begin{vmatrix}a & b\\ c & d\end{vmatrix}\neq 0$$

   则(U,V)也服从二维正态分布

1.5 随机变量的数字特征

1.5.1 期望定义

1. 定义：

   (a) 对于离散型随机变量X，设分布律为$P(X=x_k)=p_k,\ k=1,2,\cdots$，若级数$\sum\limits_{k=1}^{\infty} x_k p_k$绝对收敛，则$X$的期望定义为

   $$E(X)=\sum_{k=1}^{\infty} x_k p_k$$

   (b) 对于连续型随机变量X，设概率密度函数为$f(x)$，若积分$\int_{-\infty}^{+\infty} x f(x)\,dx$绝对收敛，则$X$的期望定义为

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

2. 一维随机变量函数期望：
   设函数$g(x)$连续，则随机变量$Y=g(X)$的期望如下

   (a) 对于离散型随机变量X，设分布律为$P(X=x_k)=p_k,\ k=1,2,\cdots$，若级数$\sum\limits_{k=1}^{\infty} g(x_k)\,p_k$绝对收敛，则有

   $$E(Y)=E[g(X)]=\sum_{k=1}^{\infty} g(x_k)\,p_k$$

   (b) 对于连续型随机变量X，设概率密度函数为$f(x)$，若积分$\int_{-\infty}^{+\infty} g(x)\,f(x)\,dx$绝对收敛，则有

   $$E(Y)=E[g(X)]=\int_{-\infty}^{+\infty} g(x)\,f(x)\,dx$$

3. 二维随机变量函数期望：
   设函数$g(x,y)$连续，则随机变量$Z=g(X,Y)$的期望如下

   (a) 对于二维离散型随机变量(X,Y)，设分布律为$P\left(X=x_i,Y=y_j\right)=p_{ij},\ i,j=1,2,\cdots$，若级数$\sum\limits_{i=1}^{\infty}\sum\limits_{j=1}^{\infty} g\left(x_i,y_j\right)p_{ij}$绝对收敛，则有

   $$E(Z)=E[g(X,Y)]=\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} g\left(x_i,y_j\right)p_{ij}$$

   特别地，我们有

   $$E(X)=\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} x_i p_{ij}=\sum_{i=1}^{\infty} x_i p_{i\cdot}$$

   $$E(Y)=\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} y_j p_{ij}=\sum_{j=1}^{\infty} y_j p_{\cdot j}$$

(b) 对于连续型随机变量(X,Y)，设概率密度函数为$f(x,y)$，若积分$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)\,f(x,y)\,dxdy$绝对收敛，则有

$$E(Z)=E[g(X,Y)]=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)\,f(x,y)\,dxdy$$

特别地，我们有

$$E(X)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xf(x,y)\,dxdy=\int_{-\infty}^{+\infty} xf_X(x)\,dx$$

$$E(Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} yf(x,y)\,dxdy=\int_{-\infty}^{+\infty} yf_Y(y)\,dy$$

1.5.2 方差定义

1. 定义：随机变量X的期望存在时若$E[X-E(X)]^2$也存在，则有

$$D(X)=E[X-E(X)]^2=E\left(X^2\right)-[E(X)]^2$$

而标准差定义为

$$\sigma(X)=\sqrt{D(X)}$$

(a) 对于离散型随机变量X，设分布律为$P(X=x_k)=p_k$，$k=1,2,\cdots$，若其期望$E(X)$存在，$\sum\limits_{k=1}^{\infty}[x_k-E(X)]^2\,p_k$绝对收敛，则$X$的方差定义为

$$D(X)=\sum_{k=1}^{\infty}[x_k-E(X)]^2\,p_k$$

(b) 对于连续型随机变量X，设概率密度函数为$f(x)$，若其期望$E(X)$存在，$\int_{-\infty}^{+\infty}[x-E(X)]^2\,f(x)\,dx$绝对收敛，则$X$的方差定义为

$$D(X)=\int_{-\infty}^{+\infty}[x-E(X)]^2\,f(x)\,dx$$

2. 一维随机变量函数方差：
设函数$g(x)$连续，则随机变量$Y=g(X)$的方差为

$$D(Y)=D[g(X)]=E\left[g(X)^2\right]-\{E[g(X)]\}^2$$

(a) 对于离散型随机变量X，设分布律为$P(X=x_k)=p_k$，$k=1,2,\cdots$，若其期望$E(X)$存在，$\sum\limits_{k=1}^{\infty}[g(x_k)-E(X)]^2\,p_k$绝对收敛，则$X$的方差定义为

$$D(X)=\sum_{k=1}^{\infty}[g(x_k)-E(X)]^2\,p_k$$

(b) 对于连续型随机变量X，设概率密度函数为$f(x)$，若其期望$E(X)$存在，$\int_{-\infty}^{+\infty}[g(x)-E(X)]^2 f(x)\,dx$绝对收敛，则$X$的方差定义为

$$D(X)=\int_{-\infty}^{+\infty}[g(x)-E(X)]^2 f(x)\,dx$$

3. 二维随机变量函数方差：
设函数$g(x,y)$连续，则随机变量$Z=g(X,Y)$的方差为

(a) 对于二维离散型随机变量(X,Y)，设分布律为$P\left(X=x_i,Y=y_j\right)=p_{ij}$，$i,j=1,2,\cdots$，若级数$\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\left\{g\left(x_i,y_j\right)-E[g(X,Y)]\right\}^2 p_{ij}$绝对收敛，则有

$$D(Z)=D[g(X,Y)]=\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\left\{g\left(x_i,y_j\right)-E[g(X,Y)]\right\}^2 p_{ij}$$

特别地，我们有

$$D(X)=\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}[x_i-E(X)]^2 p_{ij}=\sum_{i=1}^{\infty}[x_i-E(X)]^2 p_{i\cdot}$$

$$D(Y)=\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\left[y_j-E(Y)\right]^2 p_{ij}=\sum_{j=1}^{\infty}\left[y_j-E(Y)\right]^2 p_{\cdot j}$$

(b) 对于连续型随机变量(X,Y)，设概率密度函数为$f(x,y)$，若积分$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}g(x,y)f(x,y)\,dxdy$绝对收敛，则有

$$D(Z)=D[g(X,Y)]=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\{g(x,y)-E[g(X,Y)]\}^2 f(x,y)\,dxdy$$

特别地，我们有

$$D(X)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}[x-E(X)]^2 f(x,y)\,dxdy=\int_{-\infty}^{+\infty}[x-E(X)]^2 f_X(x)\,dx$$

$$D(Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}[y-E(Y)]^2 f(x,y)\,dxdy=\int_{-\infty}^{+\infty}[y-E(Y)]^2 f_Y(y)\,dy$$

1.5.3 期望与方差的性质

1. $E(kX\pm c)=kE(X)\pm c,\ D(kX\pm c)=k^2D(X)$

2. $E(X_1\pm X_2)=E(X_1)\pm E(X_2)\,,\ D(X_1\pm X_2)=D(X_1)+D(X_2)\pm 2\operatorname{Cov}(X_1,X_2)$

3. 若X与Y不相关，则有

$$E(XY)=E(X)E(Y)$$

4. 若X与Y相互独立，则有

$$D(XY)=D(X)D(Y)+[E(X)]^2D(Y)+[E(Y)]^2D(X)$$

5. 若X与Y相互独立，$g(x)$和$h(x)$是两个连续函数，则$U=g(X)$和$V=h(Y)$也都是随机变量，且U和V相互独立

1.5.4 协方差与相关系数

1. 定义：

   (a) 协方差$\mathrm{Cov}(X,Y)$定义为：

   $$\mathrm{Cov}(X,Y)=E\{[X-E(X)][Y-E(Y)]\}=E(XY)-E(X)E(Y)$$

      i. 二维离散型随机变量协方差：

      $$\mathrm{Cov}(X,Y)=\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}[x_i-E(X)]\left[y_j-E(Y)\right]p_{ij}$$

      ii. 二维连续型随机变量协方差：

      $$\mathrm{Cov}(X,Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}[x-E(X)][y-E(Y)]f(x,y)\,dxdy$$

   (b) 相关系数ρ_{XY}定义为：

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

2. 协方差性质：

   (a) $\mathrm{Cov}(X,Y)=\mathrm{Cov}(Y,X)$

   (b) $\mathrm{Cov}(X,k)=0$，$\mathrm{Cov}(k,Y)=0$

   (c) $\mathrm{Cov}(aX\pm b,cY\pm d)=ac\,\mathrm{Cov}(X,Y)$

   (d) $\mathrm{Cov}(X_1\pm X_2,Y)=\mathrm{Cov}(X_1,Y)\pm\mathrm{Cov}(X_2,Y)$

3. 相关系数性质：

   (a) $|\rho_{XY}|\le 1$

   (b) $|\rho_{XY}|=1$的充要条件是，存在常数a,b使

   $$P\{Y=aX+b\}=1$$

   成立（即相关系数描述的是随机变量的线性相关性，$\rho_{XY}=0$时即表示X,Y不相关）

1.5.5 矩与协方差矩阵

1. 矩定义：

   (a) k阶原点矩：$E\left(X^k\right)$

   (b) k阶中心矩：$E\left\{[X-E(X)]^k\right\}$

   (c) $k+l$阶混合矩：$E\left(X^kY^l\right)$

(d) $k+l$阶混合中心矩：$E\left\{[X-E(X)]^k[Y-E(Y)]^l\right\}$

2. **协方差矩阵定义**：设n维随机变量$(X_1, X_2, \cdots, X_n)$的协方差

$$c_{ij} = c_{ji} = \mathrm{Cov}\left(X_i, X_j\right)$$

均存在，则称矩阵

$$C = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix}$$

为n维随机变量$(X_1, X_2, \cdots, X_n)$的协方差矩阵，显然协方差矩阵为一个对称矩阵

1.5.6 其他

1. **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，我们有下述不等式成立：

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

1.6 大数定律与中心极限定理

1.6.1 依概率收敛

1. **定义**：设$X_1, X_2, \cdots$是一个随机变量序列，a是一个常数，若对于任意$\varepsilon > 0$，有

$$\lim_{n\to\infty} P\{|X_n - a| < \varepsilon\} = 1$$

则称序列$X_1, X_2, \cdots$依概率收敛于a，记为

$$X_n \xrightarrow{P} a$$

2. **性质**：设$X_n \xrightarrow{P} a$，$Y_n \xrightarrow{P} b$，又函数$g(x, y)$在点(a, b)处连续，则

$$g(X_n, Y_n) \xrightarrow{P} g(a, b)$$

1.6.2 大数定律

1. **大数定律**：设$X_1, X_2, \cdots$为一随机变量序列，若对于任意$\varepsilon > 0$，有

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{k=1}^{n} X_k - E\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right)\right| < \varepsilon\right\} = 1$$

则称随机变量序列$X_1, X_2, \cdots$服从大数定律

2. **切比雪夫大数定律**：设$X_1, X_2, \cdots$是相互独立的随机变量序列，且方差一致有上界，则对于任意$\varepsilon > 0$，有

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \frac{1}{n}\sum_{k=1}^{n} E(X_k)\right| < \varepsilon\right\} = 1$$

3. **辛钦大数定律（弱大数定律）**：设$X_1, X_2, \cdots$是相互独立且服从同一分布的随机变量序列，且具有数学期望$E(X_k) = \mu$，则对于任意$\varepsilon > 0$，有

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \mu\right| < \varepsilon\right\} = 1$$

   即$\overline{X} = \frac{1}{n}\sum\limits_{k=1}^{n} X_k$依概率收敛于$\mu$，$\overline{X} \xrightarrow{P} \mu$

4. **伯努利大数定律**：设n_A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$\varepsilon > 0$，有

$$\lim_{n\to\infty} P\left\{\left|\frac{n_A}{n} - p\right| < \varepsilon\right\} = 1$$

   即事件A发生的频率依概率收敛于事件A发生的概率，$\dfrac{n_A}{n} \xrightarrow{P} p$

1.6.3 中心极限定理

1. **李雅普诺夫中心极限定理**：设随机变量序列$X_1, X_2, \cdots$相互独立，且具有期望和方差

$$E(X_k) = \mu_k,\ D(X_k) = \sigma_k^2 > 0$$

   则随机变量

$$Z_n = \frac{\sum\limits_{k=1}^{n} X_k - \sum\limits_{k=1}^{n} \mu_k}{\sqrt{\sum\limits_{k=1}^{n} \sigma_k^2}}$$

   的分布函数$F_n(x)$对任意实数x满足

$$\lim_{n\to\infty} F_n(x) = \lim_{n\to\infty} P\left\{\frac{\sum\limits_{k=1}^{n} X_k - \sum\limits_{k=1}^{n} \mu_k}{\sqrt{\sum\limits_{k=1}^{n} \sigma_k^2}} \le x\right\} = \Phi(x)$$

   即当n很大时，随机变量Z_n近似地服从标准正态分布$N(0,1)$

2. **列维-林德伯格中心极限定理（独立同分布的中心极限定理）**：设随机变量序列$X_1, X_2, \cdots$相互独立且同分布，且具有期望和方差

$$E(X_k) = \mu,\ D(X_k) = \sigma^2 > 0$$

则随机变量

$$Y_n = \frac{\sum\limits_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma}$$

的分布函数$F_n(x)$对任意实数x满足

$$\lim_{n\to\infty} F_n(x) = \lim_{n\to\infty} P\left\{\frac{\sum\limits_{k=1}^{n} X_k - n\mu}{\sqrt{n}\sigma} \leq x\right\} = \Phi(x)$$

即当n很大时，随机变量Y_n近似地服从标准正态分布$N(0,1)$，或者说，当n很大时随机变量序列的算术平均$\overline{X} = \frac{1}{n}\sum\limits_{k=1}^{n} X_k$服从均值为$\mu$方差为$\frac{\sigma^2}{n}$的正态分布

3. **棣莫弗-拉普拉斯中心极限定理**：本定理是列维-林德伯格中心极限定理的特殊情况。设伯努利试验中试验A出现的概率为p，记n重伯努利试验中A出现的次数为η_n（即η_n服从参数为n,p的二项分布），则对任意实数x有

$$\lim_{n\to\infty} P\left\{\frac{\eta_n - np}{\sqrt{np(1-p)}} \leq x\right\} = \Phi(x)$$

1.7 样本及抽样分布

1.7.1 基本定义

1. **样本**：设X是具有分布函数F的随机变量，若$X_1, X_2, \cdots, X_n$相互独立且均有同一分布函数F，则称$X_1, X_2, \cdots, X_n$为从总体X得到的容量为n的简单随机样本，简称样本，它们的观察值$x_1, x_2, \cdots, x_n$称为样本值，又称为X的n个独立的观察值

2. **统计量**：设$X_1, X_2, \cdots, X_n$是来自总体X的一个样本，$g(X_1, X_2, \cdots, X_n)$是$X_1, X_2, \cdots, X_n$的函数，若g中不含未知参数，则称$g(X_1, X_2, \cdots, X_n)$是一统计量，统计量的分布就称为抽样分布

   常见统计量有：

   (a) 样本均值：

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

   (b) 样本方差：

   $$S^2 = \frac{\sum\limits_{i=1}^{n}\left(X_i - \overline{X}\right)^2}{n-1} = \frac{\sum\limits_{i=1}^{n} X_i^2 - n\overline{X}}{n-1}$$

   (c) 样本标准差：

   $$S = \sqrt{\frac{\sum\limits_{i=1}^{n}\left(X_i - \overline{X}\right)^2}{n-1}} = \sqrt{\frac{\sum\limits_{i=1}^{n} X_i^2 - n\overline{X}}{n-1}}$$

1.7.2 Γ函数

1. Γ函数的定义为：

$$\Gamma(s) = \int_0^{+\infty} e^{-x} x^{s-1} dx \quad (s > 0)$$

2. 计算性质：

(a) $\Gamma(1) = 1$, $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

(b) $\Gamma(s+1) = s\Gamma(s)$（即我们有，对任意正整数n，$\Gamma(n+1) = n!$）

(c) $\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin \pi s}$

(d) $\lim\limits_{s \to 0^+} \Gamma(s) = +\infty$

1.7.3 抽样分布

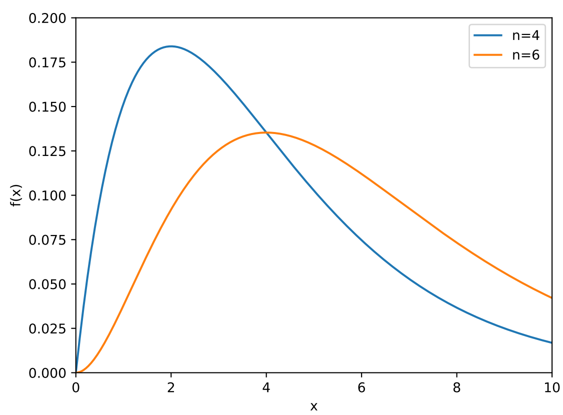
1. χ^2**分布（卡方分布）**：设$X_1, X_2, \cdots, X_n$是来自总体$N(0,1)$的样本，则称统计量

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

服从自由度为n的χ^2分布，记为$\chi^2 \sim \chi^2(n)$，概率密度函数为

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

函数图像为：



χ^2分布性质：

(a) 设$\chi_1^2 \sim \chi^2(n_1)$，$\chi_2^2 \sim \chi^2(n_2)$，且$\chi_1^2$和$\chi_2^2$相互独立，则有

$$\chi_1^2 + \chi_2^2 \sim \chi^2(n_1 + n_2)$$

(b) 设$\chi^2 \sim \chi^2(n)$，则有

$$E\left(\chi^2\right) = n,\ D\left(\chi^2\right) = 2n$$

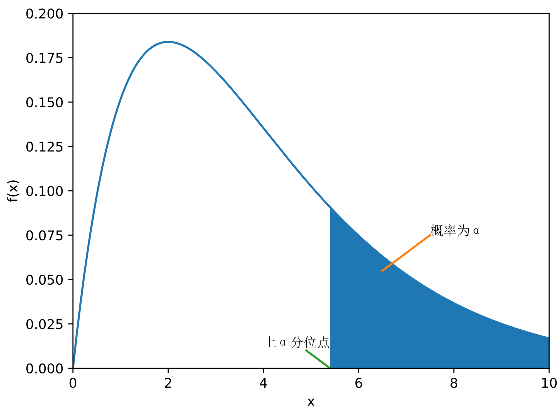
(c) 设总体$X \sim N\left(\mu, \sigma^2\right)$，$X_1, X_2, \cdots, X_n$是来自总体$X$的样本，则

$$\frac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\sigma^2} \sim \chi^2(n)$$

(d) 设$\chi^2 \sim \chi^2(n)$，对于给定的正数$\alpha \in (0,1)$，则满足

$$P\left\{\chi^2 > \chi^2_\alpha(n)\right\} = \alpha$$

的点$x = \chi^2_\alpha(n)$称为χ^2的上α分位点



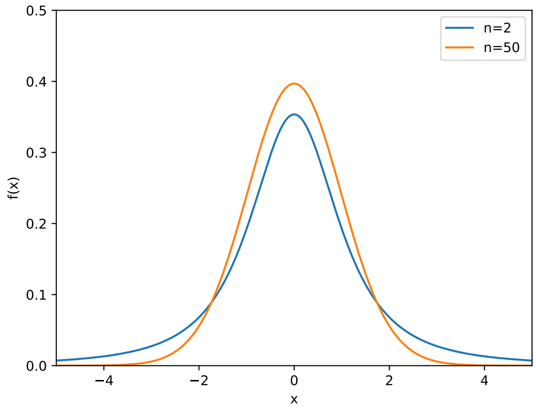
2. **t分布**：设$X \sim N(0,1)$，$Y \sim \chi^2(n)$，且X, Y相互独立，则称随机变量

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

服从自由度为n的t分布，记为$t \sim t(n)$，概率密度函数为

$$f(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\cdot\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}$$

函数图像为：



t分布性质：

(a) 对于t分布的密度函数$f(x)$我们有

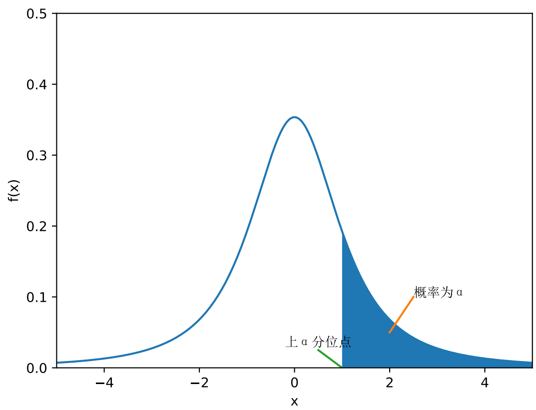
$$\lim_{n\to\infty} f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

即当n足够大时t分布近似于标准正态分布$N(0,1)$

(b) 设$t \sim t(n)$，对于给定的正数$\alpha \in (0,1)$，则满足

$$P\{t > t_\alpha(n)\} = \alpha$$

的点$x = t_\alpha(n)$称为t的上α分位点



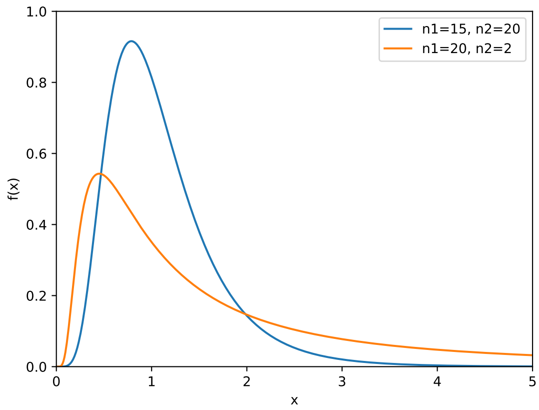
3. **F分布**：设$U \sim \chi^2(n_1)$，$V \sim \chi^2(n_2)$，且U, V相互独立，则称随机变量

$$F = \frac{\frac{U}{n_1}}{\frac{V}{n_2}}$$

服从自由度为(n_1, n_2)的F分布，记为$F \sim F(n_1, n_2)$，概率密度函数为

$$f(x) = \begin{cases} \dfrac{\Gamma\left(\dfrac{n_1+n_2}{2}\right) \cdot \left(\dfrac{n_1}{n_2}\right)^{\frac{n_1}{2}} \cdot x^{\frac{n_1}{2}-1}}{\Gamma\left(\dfrac{n_1}{2}\right) \cdot \Gamma\left(\dfrac{n_2}{2}\right) \cdot \left(1 + \dfrac{n_1}{n_2}x\right)^{\frac{n_1+n_2}{2}}}, & x > 0 \\ 0, & else \end{cases}$$

函数图像为：



F分布性质：

(a) 若$F \sim F(n_1, n_2)$，则有

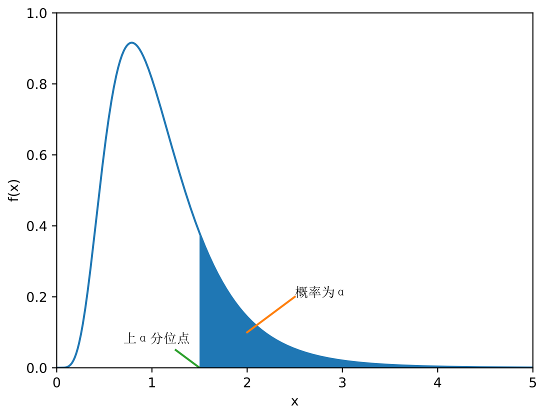
$$\frac{1}{F} \sim F(n_2, n_1)$$

(b) 若$X \sim t(n)$，则$X^2 \sim F(1, n)$

(c) 设$F \sim F(n_1, n_2)$，对于给定的正数$\alpha \in (0, 1)$，则满足

$$P\{F > F_\alpha(n_1, n_2)\} = \alpha$$

的点$x = F_\alpha(n_1, n_2)$称为F的上α分位点



1.7.4 抽样分布样本均值和方差的分布

1. 设总体X具有均值μ和方差σ^2，$X_1, X_2, \cdots, X_n$是来自X的一个样本，$\overline{X}$和S^2分别是样本均值和样本方差，则我们有

$$E\left(\overline{X}\right) = \mu,\ D\left(\overline{X}\right) = \frac{\sigma^2}{n},\ E\left(S^2\right) = \sigma^2$$

2. **单个正态总体样本均值分布**：设$X_1, X_2, \cdots, X_n$是来自正态总体$N\left(\mu, \sigma^2\right)$的一个样本，$\overline{X}$是样本均值，则有

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

3. **两个正态总体样本均值分布**：设$X_1, X_2, \cdots, X_m$与$Y_1, Y_2, \cdots, Y_n$分别是来自正态总体$N\left(\mu_1, \sigma_1^2\right)$和$N\left(\mu_2, \sigma_2^2\right)$的样本，且这两个正态总体$N\left(\mu_1, \sigma_1^2\right)$和$N\left(\mu_2, \sigma_2^2\right)$相互独立，$\overline{X}, \overline{Y}$分别是其样本均值，则有

$$\overline{X} \pm \overline{Y} \sim N\left(\mu_1 \pm \mu_2, \frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}\right)$$

4. **非正态总体样本均值近似分布**：设$X_1, X_2, \cdots, X_n$是来自任意总体X的一个样本，总体的期望为$E(X) = \mu$，方差为$D(X) = \sigma^2$，则当n较大时，有

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

5. 设$X_1, X_2, \cdots, X_n$是来自正态总体$N\left(\mu, \sigma^2\right)$的一个样本，$\overline{X}$和$S^2$分别是样本均值和样本方差，则$\overline{X}$和$S^2$相互独立，且有

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

6. 设$X_1, X_2, \cdots, X_n$是来自正态总体$N\left(\mu, \sigma^2\right)$的一个样本，$\overline{X}$和$S^2$分别是样本均值和样本方差，则有

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

7. 设$X_1, X_2, \cdots, X_{n_1}$与$Y_1, Y_2, \cdots, Y_{n_2}$分别是来自正态总体$N\left(\mu_1, \sigma_1^2\right)$和$N\left(\mu_2, \sigma_2^2\right)$的样本，且这两个正态总体$N\left(\mu_1, \sigma_1^2\right)$ 和$N\left(\mu_2, \sigma_2^2\right)$相互独立，$\overline{X}, \overline{Y}$ 和S_1^2, S_2^2分别是样本均值和样本方差，则有

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

而当$\sigma_1^2 = \sigma_2^2 = \sigma^2$时，还有

$$\frac{\left(\overline{X} - \overline{Y}\right) - (\mu_1 - \mu_2)}{\sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2)$$

1.8 参数估计

1.8.1 矩估计

1. **样本原点矩与中心矩**：对于总体X，设$X_1, X_2, \cdots, X_n$为其样本，则样本原点矩与中心矩定义为：

   (a) 样本k阶原点矩：
   $$A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$$

   (b) 样本k阶中心矩：
   $$B_k = \frac{1}{n}\sum_{i=1}^{n} \left(X_i - \overline{X}\right)^k$$

2. **矩估计定义**：表示总体原点矩或中心矩与未知参数关系的函数方程（组）
   $$\begin{cases} \theta_1 = \theta_1\left(\mu_1, \mu_2, \cdots, \mu_k\right) \\ \theta_2 = \theta_2\left(\mu_1, \mu_2, \cdots, \mu_k\right) \\ \quad \vdots \\ \theta_k = \theta_k\left(\mu_1, \mu_2, \cdots, \mu_k\right) \end{cases}$$
   称为矩估计方程（组），用对应的样本矩替代方程组中的总体矩后得到的
   $$\hat{\theta}_i = \theta_i\left(A_1, A_2, \cdots, A_k\right)$$
   分别称为θ_i的矩估计量，矩估计量的观察值称为矩估计值

3. **矩估计性质**：

   (a) 若$\hat{\theta}$为参数θ的矩估计量，$g(x)$为x的连续函数，则$g\left(\hat{\theta}\right)$是$g(\theta)$的矩估计量

   (b) 若$\hat{\theta}_1, \hat{\theta}_2$为参数$\theta_1, \theta_2$的矩估计量，$g(x, y)$为$(x, y)$的连续函数，则$g\left(\hat{\theta}_1, \hat{\theta}_2\right)$是$g(\theta_1, \theta_2)$的矩估计量

1.8.2 极大似然估计

1. **单参数似然函数**：对于总体X，设$f(x;\theta)$为总体X的概率密度函数，其中θ是未知参数，$X_1, X_2, \cdots, X_n$为总体X的一个样本，则自变量为θ，定义域为Θ的非负函数
   $$L\left(x_1, x_2, \cdots, x_n; \theta\right) = \prod_{i=1}^{n} f\left(x_i; \theta\right), \ \theta \in \Theta$$
   称为容量为n的样本似然函数，或称为样本值$x_1, x_2, \cdots, x_n$的似然函数

2. **单参数极大似然估计**：对于总体X，设$X_1, X_2, \cdots, X_n$为其样本，若存在$\hat{\theta} = \hat{\theta}(x_1, x_2, \cdots, x_n)$，使

$$L\left(\hat{\theta}\right) = \max_{\theta \in \Theta} L(\theta)$$

则称似然函数的最大值点$\hat{\theta}(x_1, x_2, \cdots, x_n)$为参数$\theta$的极大似然估计值，而称$\hat{\theta}(X_1, X_2, \cdots, X_n)$为参数$\theta$的极大似然估计量

3. **多参数似然函数**：对于总体X，设$f(x; \theta_1, \theta_2, \cdots, \theta_m)$为总体$X$的概率密度函数，其中$\theta_i$是未知参数，$X_1, X_2, \cdots, X_n$为总体$X$的一个样本，则自变量为$\theta_1, \theta_2, \cdots, \theta_m$，定义域为$\Theta^m$的非负函数

$$L(x_1, x_2, \cdots, x_n; \theta_1, \theta_2, \cdots, \theta_m) = \prod_{i=1}^{n} f(x_i; \theta_1, \theta_2, \cdots, \theta_m),\ (\theta_1, \theta_2, \cdots, \theta_m) \in \Theta^m$$

称为容量为n的样本似然函数，或称为样本值$x_1, x_2, \cdots, x_n$的似然函数

4. **多参数极大似然估计**：对于总体X，设$X_1, X_2, \cdots, X_n$为其样本，若存在一组$\hat{\theta}_i = \hat{\theta}_i(x_1, x_2, \cdots, x_m)$，使

$$L\left(\hat{\theta}_1, \hat{\theta}_2, \cdots, \hat{\theta}_m\right) = \max_{(\theta_1, \theta_2, \cdots, \theta_m) \in \Theta^m} L(\theta_1, \theta_2, \cdots, \theta_m)$$

则称似然函数的最大值点$\hat{\theta}_i(x_1, x_2, \cdots, x_n)$为参数$\theta_i$的极大似然估计值，而称$\hat{\theta}_i(X_1, X_2, \cdots, X_n)$为参数$\theta_i$的极大似然估计量

5. **极大似然估计性质**：

   (a) 若$\hat{\theta}$为参数θ的极大似然估计量，$g(x)$为x的连续函数，则$g\left(\hat{\theta}\right)$是$g(\theta)$的极大似然估计量

   (b) 若$\hat{\theta}_1, \hat{\theta}_2$为参数$\theta_1, \theta_2$的极大似然估计量，$g(x, y)$为$(x, y)$的连续函数，则$g\left(\hat{\theta}_1, \hat{\theta}_2\right)$是$g(\theta_1, \theta_2)$的极大似然估计量

   (c) 若任意总体X的数学期望和方差均存在，则其数学期望的矩估计量和极大似然估计量相同，均为$\frac{1}{n}\sum\limits_{i=1}^{n} X_i$；其方差的矩估计量和极大似然估计量也相同，均为$\frac{1}{n}\sum\limits_{i=1}^{n}\left(X_i - \overline{X}\right)^2$

1.8.3 估计量无偏性、有效性、一致性

1. **无偏性**：设$\hat{\theta}$是未知参数θ的估计量，若$E\left(\hat{\theta}\right) = \theta$，则称$\hat{\theta}$是$\theta$的无偏估计量

2. **有效性**：设$\hat{\theta}_1(X_1, X_2, \cdots, X_n)$与$\hat{\theta}_2(X_1, X_2, \cdots, X_n)$均为$\theta$的无偏估计量，若有

$$D\left(\hat{\theta}_1\right) < D\left(\hat{\theta}_2\right)$$

则称$\hat{\theta}_1$比$\hat{\theta}_2$更有效

3. **一致性（相合性）**：设$\hat{\theta}_n(X_1, X_2, \cdots, X_n)$为参数$\theta$的估计量，若对任意$\theta \in \Theta$，当样本容量$n \to \infty$时，$\hat{\theta}_n(X_1, X_2, \cdots, X_n)$依概率收敛于$\theta$，即对任意$\varepsilon > 0$，有

$$\lim_{n\to\infty} P\left(\left|\hat{\theta}_n - \theta\right| \geq \varepsilon\right) = 0$$

则称$\hat{\theta}_n$为θ一致估计量，也叫相合估计量

4. **性质**：
   (a) 设$\hat{\theta}_n$为θ无偏估计量，且$\lim\limits_{n\to\infty} D\left(\hat{\theta}_n\right) = 0$，则$\hat{\theta}_n$为$\theta$一致估计量
   (b) 设$\hat{\theta}_n$为θ的一致估计量，若函数$g(x)$连续，则$g\left(\hat{\theta}_n\right)$为$g(\theta)$的一致估计量
   (c) 任意样本原点矩是总体原点矩的一致估计量。进一步地，若待估参数$\theta = g(\mu_1, \mu_2, \cdots, \mu_k)$，而函数$g$连续，则矩估计量$\hat{\theta} = g(A_1, A_2, \cdots, A_k)$为$\theta$的一致估计量

1.8.4 正态总体参数的区间估计

1. 一个正态总体
   (a) **待估参数为μ，且σ^2已知时**：
   用作估计的随机变量及其分布为：

$$Z = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1)$$

   置信度为$1-\alpha$的置信区间为：

$$\left(\overline{X} \pm \frac{\sigma}{\sqrt{n}} z_{\frac{\alpha}{2}}\right)$$

   单侧置信限为：

$$\underline{\mu} = \overline{X} - \frac{\sigma}{\sqrt{n}} z_\alpha$$
$$\overline{\mu} = \overline{X} + \frac{\sigma}{\sqrt{n}} z_\alpha$$

   (b) **待估参数为μ，且σ^2未知时**：
   用作估计的随机变量及其分布为：

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

   置信度为$1-\alpha$的置信区间为：

$$\left(\overline{X} \pm \frac{S}{\sqrt{n}} t_{\frac{\alpha}{2}}(n-1)\right)$$

   单侧置信限为：

$$\underline{\mu} = \overline{X} - \frac{S}{\sqrt{n}} t_\alpha(n-1)$$
$$\overline{\mu} = \overline{X} + \frac{S}{\sqrt{n}} t_\alpha(n-1)$$

(c) **待估参数为σ^2，且μ未知时：**
用作估计的随机变量及其分布为：

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

置信度为$1-\alpha$的置信区间为：

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

单侧置信限为：

$$\underline{\sigma^2} = \frac{(n-1)\,S^2}{\chi^2_{\alpha}\,(n-1)}$$

$$\overline{\sigma^2} = \frac{(n-1)\,S^2}{\chi^2_{1-\alpha}\,(n-1)}$$

2. 两个正态总体

(a) **待估参数为$\mu_1-\mu_2$，且σ_1^2, σ_2^2已知时：**
用作估计的随机变量及其分布为：

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

置信度为$1-\alpha$的置信区间为：

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

单侧置信限为：

$$\underline{\mu_1-\mu_2} = \left(\overline{X}-\overline{Y}\right) - \sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}} \cdot z_{\alpha}$$

$$\overline{\mu_1-\mu_2} = \left(\overline{X}-\overline{Y}\right) + \sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}} \cdot z_{\alpha}$$

(b) **待估参数为$\mu_1-\mu_2$，且$\sigma_1^2=\sigma_2^2=\sigma^2$未知时：**
用作估计的随机变量及其分布为：

$$\frac{\left(\overline{X}-\overline{Y}\right)-(\mu_1-\mu_2)}{\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}} \cdot \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t\,(n_1+n_2-2)$$

置信度为$1-\alpha$的置信区间为：

$$\left(\left(\overline{X}-\overline{Y}\right) \pm \sqrt{\frac{(n_1-1)\,S_1^2+(n_2-1)\,S_2^2}{n_1+n_2-2}} \cdot \sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \cdot t_{\frac{\alpha}{2}}\,(n_1+n_2-2)\right)$$

单侧置信限为：

$$\underline{\mu_1-\mu_2}=\left(\overline{X}-\overline{Y}\right)-\sqrt{\frac{(n_1-1)\,S_1^2+(n_2-1)\,S_2^2}{n_1+n_2-2}}\cdot\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\cdot t_\alpha\,(n_1+n_2-2)$$

$$\overline{\mu_1-\mu_2}=\left(\overline{X}-\overline{Y}\right)+\sqrt{\frac{(n_1-1)\,S_1^2+(n_2-1)\,S_2^2}{n_1+n_2-2}}\cdot\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\cdot t_\alpha\,(n_1+n_2-2)$$

(c) **待估参数为$\frac{\sigma_1^2}{\sigma_2^2}$，且$\mu_1,\mu_2$未知时：**
用作估计的随机变量及其分布为：

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

置信度为$1-\alpha$的置信区间为：

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

单侧置信限为：

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

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

3. 一些性质

(a) 若$\left[\underline{\theta},\overline{\theta}\right]$为$\theta$的置信度为$1-\alpha$的置信区间，$g\,(x)$为单调函数，则$\left[g\left(\underline{\theta}\right),g\left(\overline{\theta}\right)\right]$（或$\left[g\left(\overline{\theta}\right),g\left(\underline{\theta}\right)\right]$）为$g\,(\theta)$的置信度为$1-\alpha$的置信区间

Chapter 2

解题笔记

2.1 多维随机变量及其分布

1. 离散型随机变量和连续型随机变量混合

   例：设随机变量X与Y独立，其中X的概率分布为

   $$X \sim \begin{bmatrix} 1 & 2 \\ 0.3 & 0.7 \end{bmatrix}$$

   而Y的概率密度为$f(Y)$，求随机变量$U = X + Y$的概率密度$g(u)$

   解：设$F(y)$是Y的分布函数，则由全概公式以及X和Y相互独立知$U = X + Y$的分布函数为

   $$\begin{aligned} G(u) &= P(X + Y \leq u) \\ &= 0.3P(X + Y \leq u|X = 1) + 0.7P(X + Y \leq u|X = 2) \\ &= 0.3P(Y \leq u - 1|X = 1) + 0.7P(Y \leq u - 2|X = 2) \\ &= 0.3P(Y \leq u - 1) + 0.7P(Y \leq u - 2) \\ &= 0.3F(u - 1) + 0.7F(u - 2) \end{aligned}$$

   因此$U = X + Y$的概率密度为

   $$g(u) = G'(u) = 0.3f(u - 1) + 0.7f(u - 2)$$

2.2 随机变量的数字特征

1. 利用随机变量分解法和数字特征运算规律求解数字特征

   例：将n个球（$1 \sim n$号）随机放进n个盒子（$1 \sim n$号）中去，一个盒子装一个球，将一个球装入与球同号的盒子中称为一个配对，记X为总的配对数，求$E(X)$

   解：记随机变量X_i为i号球的配对数，显然每个球的配对数只有1和0两种可能，而尽管

随机变量X_i互相并不独立，但我们有$P(X_i = 1) = \frac{1}{n}$，而又因为$X = \sum\limits_{i=1}^{n} X_i$，故我们有

$$E(X) = \sum_{i=1}^{n} E(X_i) = n \cdot \frac{1}{n} = 1$$

2.3 样本及抽样分布

1. 利用χ^2分布的期望和方差易求的特性处理统计量期望或者方差的求解问题
   例：设$X_1, X_2, \cdots, X_n$是来自总体$X \sim N\left(\mu, \sigma^2\right)$的简单随机样本，样本均值为$\overline{X} = \frac{1}{n}\sum\limits_{i=1}^{n} X_i$，记$S_0^2 = \frac{1}{n}\sum\limits_{i=1}^{n}(X_i-\mu)^2$，$S_1^2 = \frac{1}{n}\sum\limits_{i=1}^{n}\left(X_i-\overline{X}\right)^2$，$S^2 = \frac{1}{n-1}\sum\limits_{i=1}^{n}\left(X_i-\overline{X}\right)^2$，当样本量$n > 2$时，求$D\left(S^2\right)$、$D\left(S_0^2\right)$、$D\left(S_1^2\right)$的大小顺序
   解：易知：
$$\frac{n}{\sigma^2}S_0^2 \sim \chi^2(n)$$
$$\frac{n}{\sigma^2}S_1^2 \sim \chi^2(n-1)$$
$$\frac{n-1}{\sigma^2}S^2 \sim \chi^2(n-1)$$
   故我们有：
$$D\left(\frac{n}{\sigma^2}S_0^2\right) = 2n \Rightarrow D\left(S_0^2\right) = \frac{2\sigma^4}{n}$$
$$D\left(\frac{n}{\sigma^2}S_1^2\right) = 2(n-1) \Rightarrow D\left(S_1^2\right) = \frac{2\sigma^4(n-1)}{n^2}$$
$$D\left(\frac{n-1}{\sigma^2}S^2\right) = 2(n-1) \Rightarrow D\left(S^2\right) = \frac{2\sigma^4}{n-1}$$
   因此，大小顺序为$D\left(S^2\right) > D\left(S_0^2\right) > D\left(S_1^2\right)$