大数定律中心极限定理

一、依概率收敛 $\left(X_n \xrightarrow{p} X\right)$

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

【注】狗 $\xrightarrow{p} a \Leftrightarrow \lim\limits_{n\to\infty} P\{|\text{狗} - a| < \varepsilon\} = 1$

【会证明依概率收敛】如：已知 $X_n \sim f_n(x) = \dfrac{n}{\pi(1+n^2x^2)}$，证明 $X_n \xrightarrow{p} 0$.

【分析】证明 $X_n \xrightarrow{p} 0 \Leftrightarrow$ 证明 $\lim\limits_{n\to\infty} P\{|X_n - 0| < \varepsilon\} = 1$，其中 $P\{|X_n - 0| < \varepsilon\} = P\{-\varepsilon < X_n < \varepsilon\} = \int_{-\varepsilon}^{\varepsilon} \dfrac{n}{\pi(1+n^2x^2)}dx = \dfrac{2n}{\pi}\int_0^{\varepsilon} \dfrac{1}{1+n^2x^2}dx$

$= \lim\limits_{n\to\infty} \dfrac{2}{\pi}\arctan n\varepsilon = 1$.证毕.

二、大数定律

$$\boxed{\begin{cases}\text{切比雪夫：①独立；②方差有界} \\ \text{辛钦：①独立同分布；②期望存在}\end{cases} \xrightarrow[\text{那就是 均值依概率收敛到均值的期望}]{\text{都在说明一件事}} \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \Leftrightarrow \lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - E\left(\frac{1}{n}\sum_{i=1}^{1} X_i\right)\right| < \varepsilon\right\} = 1}$$

【注】$\dfrac{1}{n}\sum\limits_{i=1}^{n} \text{狗}_i \xrightarrow{p} E\left(\dfrac{1}{n}\sum\limits_{i=1}^{n} \text{狗}_i\right)$

【举例】设 $X \sim E(2)$，则 $n \to \infty$ 时，$Y_n = \dfrac{1}{n}\sum\limits_{i=1}^{n} X_i^2$ 依概率收敛到______.

【分析】$X \sim E(2) \Rightarrow EX = \dfrac{1}{2}$，$DX = \dfrac{1}{4}$，因此有独立同分布，期望方差存在 $\Rightarrow \begin{cases}\text{考虑辛钦大数定律，若是做依概率收敛的题，用之} \\ \text{考虑中心极限定理，若题目中有} \begin{cases}\Sigma X_i \\ \Phi(x)\end{cases}\text{，用之（标志明显，不难看出）}\end{cases}$

显然此题用辛钦大数定律的依概率收敛功能：配合【注】中的广义化有：$\dfrac{1}{n}\sum\limits_{i=1}^{n} X_i^2 \xrightarrow{p} E\left(\dfrac{1}{n}\sum\limits_{i=1}^{n} X_i^2\right)$，而 $E\left(\dfrac{1}{n}\sum\limits_{i=1}^{n} X_i^2\right) = \dfrac{1}{n}E\left(\sum\limits_{i=1}^{n} X_i^2\right) = \dfrac{1}{n}\cdot n \cdot E(X^2)$

$= E(X^2) = DX + (EX)^2 = \dfrac{1}{4} + \left(\dfrac{1}{2}\right)^2 = \dfrac{1}{2}$. 故填 $\dfrac{1}{2}$.

三、中心极限定理

$$X_i\text{独立同分布、期望方差存在} \xRightarrow{n\to\infty} \sum_{i=1}^{n} X_i \sim N(n\mu, n\sigma^2) \xRightarrow{\text{标准化}} \frac{\sum\limits_{i=1}^{n} X_i - n\mu}{\sqrt{n}\,\sigma} \sim N(0,1)\text{，即} \lim_{n\to\infty} P\left\{\frac{\sum\limits_{i=1}^{n} X_i - n\mu}{\sqrt{n}\,\sigma} \leqslant x\right\} = \Phi(x)$$

上述为推导过程，了解即可，关于中心极限定理只需记：$\boxed{X_i\text{独立同分布，期望方差存在} \Rightarrow \lim\limits_{n\to\infty} P\left\{\dfrac{\sum\limits_{i=1}^{n} X_i - n\mu}{\sqrt{n}\,\sigma} \leqslant x\right\} = \Phi(x)}$

【注】二项分布求概率的三种方法：$X \sim B(n,p)$

①试验次数少 $\Rightarrow P\{X = k\} = C_n^k p^k (1-p)^{n-k}$

②试验次数多 $\xRightarrow{\text{泊松近似}} P\{X = k\} = \dfrac{\lambda^k}{k!}e^{-\lambda}$，其中 $\lambda = np$.

③试验次数极多 $\xRightarrow[\substack{\text{[答]实际上这里的}X\text{就是求和} \\ \text{二项分布可以看做}n\text{个}01\text{分布} \\ \text{求和得到且其和}X\sim B(n,p)}]{\substack{\text{由中心极限定理} \\ \text{[问]若用此定理不是要求和吗}}} P\{a < X < b\} = P\left\{\dfrac{a-np}{\sqrt{np(1-p)}} < \dfrac{X-np}{\sqrt{np(1-p)}} < \dfrac{b-np}{\sqrt{np(1-p)}}\right\} = \Phi\left(\dfrac{b-np}{\sqrt{np(1-p)}}\right) - \Phi\left(\dfrac{a-np}{\sqrt{np(1-p)}}\right)$.

【练习】设来自总体的样本 X_i，满足 $EX_n = 0$　(1)求 $\lim\limits_{n\to\infty} P\left\{\sum\limits_{i=1}^{n} X_i < n\right\}$　(2)若 $DX_n = \sigma^2$ 存在，求 $\lim\limits_{n\to\infty} P\left\{\sum\limits_{i=1}^{n} X_i \leqslant 0\right\}$

【分析】(1) X_i 独立同分布，期望存在 $\Rightarrow$ 辛钦大数定律：$\dfrac{1}{n}\sum\limits_{i=1}^{n} X_i \xrightarrow{p} E\left(\dfrac{1}{n}\sum\limits_{i=1}^{n} X_i\right) = \dfrac{1}{n}\sum\limits_{i=1}^{n} EX_i = 0 \xRightarrow[\text{功能}]{\text{辛钦的}} \lim\limits_{n\to\infty} P\left\{\left|\dfrac{1}{n}\sum\limits_{i=1}^{n} X_i - 0\right| < \varepsilon\right\} = 1$

又 $P\left\{\left|\dfrac{1}{n}\sum\limits_{i=1}^{n} X_i - 0\right| < \varepsilon\right\} \leqslant P\left\{\dfrac{1}{n}\sum\limits_{i=1}^{n} X_i < \varepsilon\right\} \leqslant 1 \Rightarrow n \to \infty$ 由夹逼准则 $\Rightarrow \lim\limits_{n\to\infty} P\left\{\dfrac{1}{n}\sum\limits_{i=1}^{n} X_i < \varepsilon\right\} = 1$，取 $\varepsilon = 1 \Rightarrow \lim\limits_{n\to\infty} P\left\{\sum\limits_{i=1}^{n} X_i < n\right\} = 1$.

(2)现在可以用中心极限定理了：$\lim\limits_{n\to\infty} P\left\{\dfrac{\Sigma X_i - 0}{\sqrt{n}\,\sigma} \leqslant x\right\} = \Phi(x)$，令 $x = 0 \Rightarrow \lim\limits_{n\to\infty} P\left\{\sum\limits_{i=1}^{n} X_i \leqslant 0\right\} = \Phi(0) = \dfrac{1}{2}$.

统计量及其分布

一、常用统计量

①均值：$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$　②方差：$S^2=\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2 \Rightarrow$ 标准差$S=\sqrt{S^2}$

③k阶原点矩：$A_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k$　④k阶中心矩：$B_k=\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^k$

二、四大分布

①正太分布：$X\sim N(0,1)$，$\Phi(-x)=1-\Phi(x)$

②χ^2分布：标准正太加起来：$X_i\sim N(0,1) \Rightarrow \boxed{X\sim X_1^2+X_2^2+\cdots+X_n^2\sim\chi^2(n)}$

性质：$X\sim\chi^2(n)$，$Y\sim\chi^2(m) \Rightarrow X+Y\sim\chi^2(n+m)$；$EX=n$，$DX=2n$，图像：甩鞭子，

③t分布：上面是一个标准正太，下面是卡方除以个数开根号：$X\sim N(0,1)$，$Y\sim\chi^2(n)$，$\boxed{t=\frac{X}{\sqrt{\frac{Y}{n}}}\sim t(n)}$，图像：正太分布图；

再了解一个性质：$t_{1-\alpha}(n)=-t_{\alpha}(n)$

④F分布：上面是自由度是n_1的卡方，下面是自由度是n_2的卡方：$X\sim\chi^2(n_1)$，$Y\sim\chi^2(n_2)$，$\boxed{F=\frac{\frac{X}{m}}{\frac{Y}{n}}\sim F(m,n)}$，图像：甩鞭子；

再了解一个性质：$F_{1-\alpha}(n_1,n_2)=\frac{1}{F_{\alpha}(n_2,n_1)}$

三、正太总体下的常用结论

1、X_1，$X_2\cdots X_n$来自正太总体，$X_i\sim N(\mu,\sigma^2)$，X_i服从正太分布，则$\overline{X}$(样本均值)，S^2(样本方差)独立，且有如下结论：

①对样本均值处理：$\overline{X}\sim N\left(\mu,\frac{\sigma^2}{n}\right)\xRightarrow{\text{标准化}}\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\sim N(0,1)$，叫总体的均值服从期望为$\mu$，方差为$\frac{\sigma^2}{n}$的正太分布，标准化减去$\mu$除以$\frac{\sqrt{\sigma}}{n}$服从$N(0,1)$

$$\frac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}}\sim t(n-1)$$，把上面标准化后的$\frac{\sigma}{\sqrt{n}}$，换成$\frac{S}{\sqrt{n}}$，就服从自由度少一个的t分布

②对单个样本X_i处理：$\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2\sim\chi^2(n)$，叫把一个样本减去期望$\mu$除以标准差$\sigma$标准化$\sim N(0,1)$，求和显然服从卡方分布。

$$\sum_{i=1}^{n}\left(\frac{X_i-\overline{X}}{\sigma}\right)^2\sim\chi^2(n-1)$$，把对一个样本X_i减去的期望换成均值，卡方少一个自由度，且$\sum_{i=1}^{n}\left(\frac{X_i-\overline{X}}{\sigma}\right)^2=\frac{(n-1)S^2}{\sigma^2}$

2、对于任何$EX=\mu$，$DX=\sigma^2$的总体，都有：均值的$\begin{matrix}\text{期望}\\ \text{方差}\end{matrix}\begin{cases}E\overline{X}=\mu\\ D\overline{X}=\frac{\sigma^2}{n}\end{cases}$　样本方差的$\begin{matrix}\text{期望}\\ \text{方差}\end{matrix}\begin{cases}ES^2=\sigma^2\\ DS^2=\frac{2\sigma^4}{n-1}\end{cases}$

【总结一下】 对于统计量，首先记住常用的统计量，尤其是样本方差S^2是如何定义的，然后记住统计量的四大分布，然后记住常用结论，这一节就完毕.

统计量的四大分布：①$X_i\sim N(0,1)$，②$X_1^2+X_2^2+\cdots+X_n^2\sim\chi^2(n)$，③$\frac{X_1}{\sqrt{\frac{Y_1^2+Y_2^2+\cdots+Y_n^2}{n}}}\sim t(n)$，④$\frac{X_1^2+X_2^2+\cdots+X_m^2/m}{Y_1^2+Y_2^2+\cdots+Y_n^2/n}\sim F(m,n)$ (XY独立)

四大抽样分布结论：$\begin{matrix}X_i\text{来自正太总体}\\ N(\mu,\sigma^2)\end{matrix}\Rightarrow\begin{cases}\text{①对}X_i\text{：}\frac{X_i-\mu}{\sigma}\sim N(0,1)\Rightarrow\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2\sim\chi^2(n)\Rightarrow\sum_{i=1}^{n}\left(\frac{X_1-\overline{X}}{\sigma}\right)^2\sim\chi^2(n-1)=\frac{(n-1)S^2}{\sigma^2}\\ \text{②对}\overline{X}\text{：}\overline{X}\sim N\left(\mu,\frac{\sigma^2}{n}\right)\Rightarrow\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\sim N(0,1)\Rightarrow\frac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}}\sim t(n-1)\Rightarrow\frac{\left(\overline{X}-\mu\right)^2}{\frac{S^2}{n}}\sim F(1,n-1)\end{cases}$

【实战练习】

1、X_1,X_2,X_3来自正太总体$N(0,\sigma^2)$，则统计量$S=\dfrac{X_1-X_2}{\sqrt{2}|X_3|}$服从______.

【分析】第一个题写详细一点：下面就很快了，看分子服从什么分布，$E(X_1-X_2)=EX_1-EX_2=0$，$D(X_1-X_2)\xlongequal{X_i\text{独立同分布}}DX_1+DX_2=2\sigma^2$

$\Rightarrow X_1-X_2\sim N(0,2\sigma^2)$，再看分母，见到绝对值扔进去，$\sqrt{2X_3^2}$，因为$X_3\sim N(0,\sigma^2)\Rightarrow\dfrac{X_3-0}{\sigma}\sim N(0,1)\Rightarrow\left(\dfrac{X_3}{\sigma}\right)^2\sim\chi^2(1)$

两个结合一下：分子标准化搞成标准正太和分母的卡方去凑t分布：$\dfrac{\frac{X_1-X_2-0}{\sqrt{2}\sigma}}{\sqrt{\left(\frac{X_3}{\sigma}\right)^2/1}}\sim t(1)$，整理一下：$\dfrac{X_1-X_2}{\sqrt{2}|X_3|}\sim t(1)$.

【注】对于这种求统计量分布的问题，不要沉浸在他给的形式上，你要主动出击，自己去构造，最终会发现你和出题老师的思路正好对上。

2、X_1，X_2，X_3，X_4，X_5来自正太总体$N(1,\sigma^2)$，则统计量$X=\dfrac{X_1-X_2}{|X_3+X_4-2|}$服从______.

【分析】$X_1-X_2\sim N(0,2\sigma^2),X_3+X_4\sim N(2,2\sigma^2)$,标准化：$\dfrac{X_1-X_2-0}{\sqrt{2}\sigma}\sim N(0,1)$，$\dfrac{X_3+X_4-2}{\sqrt{2}\sigma}\sim N(0,1)\Rightarrow\left(\dfrac{X_3+X_4-2}{\sqrt{2}\sigma}\right)^2\sim\chi^2(1)$

$\Rightarrow\dfrac{\frac{X_1-X_2-0}{\sqrt{2}\sigma}}{\sqrt{\left(\frac{X_3+X_4-2}{\sqrt{2}\sigma}\right)^2/1}}\sim t(1)$，且$\dfrac{\frac{X_1-X_2-0}{\sqrt{2}\sigma}}{\sqrt{\left(\frac{X_3+X_4-2}{\sqrt{2}\sigma}\right)^2/1}}=\dfrac{X_1-X_2}{|X_3+X_4-2|}\sim t(1)$.

【注】当然也可以直接写出$X_3+X_4-2\sim N(0,\sigma^2)$，因为$E(X_3+X_4-2)=EX_3+EX_4-2=0$，$D(X_3+X_4-2)=DX_3+DX_4+0=2\sigma^2$.

3、设X_1，X_2，X_3，X_4来自正太总体$N(0,2^2)$，现在有统计量$X=a(X_1-2X_2)^2+b(3X_3-4X_4)^2$，则当$a=$___$,b=$___时$,X\sim\chi^2(n),$且$n=$___.

【分析】$E(X_1-2X_2)=0$，$D(X_1-2X_2)=DX_1+4DX_2=4+16=20\Rightarrow(X_1-2X_2)\sim N(0,20)\Rightarrow\dfrac{(X_1-2X_2)-0}{\sqrt{20}}\sim N(0,1)$

$E(3X_3-4X_4)=0$，$D(3X_3-4X_4)=9DX_3+16DX_4=9\times4+16\times4=100\Rightarrow(3X_3-4X_4)\sim N(0,100)\Rightarrow\dfrac{(3X_3-4X_4)-0}{\sqrt{100}}\sim N(0,1)$

$\left(\dfrac{(X_1-2X_2)-0}{\sqrt{20}}\right)^2\sim\chi^2(1)$，$\left(\dfrac{(3X_3-4X_4)-0}{\sqrt{100}}\right)^2\sim\chi^2(1)\Rightarrow X=\left(\dfrac{(X_1-2X_2)-0}{\sqrt{20}}\right)^2+\left(\dfrac{(3X_3-4X_4)-0}{\sqrt{100}}\right)^2\sim\chi^2(2)$，且$a=\dfrac{1}{20},b=\dfrac{1}{100}$.

4、X_1，X_2，$\cdots X_n$来自总体$X\sim N(0,\sigma^2)$，则统计量$Y=\dfrac{(n-1)X_1^2}{\sum\limits_{i=2}^{n}X_i^2}$服从____.

【分析】$\dfrac{X_i-0}{\sigma}\sim N(0,1)\Rightarrow\left(\dfrac{X_i-0}{\sigma}\right)^2\sim\chi^2(1)$，$\sum\limits_{i=2}^{n}\left(\dfrac{X_i-0}{\sigma}\right)^2\sim\chi^2(n-1)$，即$\sum\limits_{i=2}^{n}\dfrac{X_i^2}{\sigma^2}\sim\chi^2(n-1)$

$\Rightarrow\dfrac{\left(\frac{X_i-0}{\sigma}\right)^2/1}{\left(\sum\limits_{i=2}^{n}\frac{X_i^2}{\sigma^2}\right)/(n-1)}\sim F(1,n-1)$，即$\dfrac{(n-1)X_i^2}{\sum\limits_{i=2}^{n}X_i^2}\sim F(1,n-1)$.

5、设随机变量X，Y都服从$N(0,1)$则
$$\begin{cases}(A)X+Y\text{服从正太分布}(\times)\text{，因为}X,Y\sim N,\text{只有独立,才有}(X,Y)\sim N\text{，进而有线性组合}\sim N\\(B)X^2+Y^2\text{服从卡方分布}(\times)\text{，还是因为没说独立，只有独立的}X_1\text{，}X_2\sim N(0,1)\text{，才有}X_1^2+X_2^2\sim\chi^2(2)\\(C)X^2,Y^2\text{都服从卡方}(\checkmark)\text{，这个没问题，}X\sim N(0,1)\Rightarrow X^2\sim\chi^2(1)\\(D)\dfrac{X^2}{Y^2}\text{服从}F\text{分布}(\times)\text{，只有}X_1\sim\chi^2(1)\text{，}X_2\sim\chi^2(1)\text{，只有}X_1X_2\text{独立才有，}\dfrac{X_1^2}{X_2^2}\sim F(1,1)\end{cases}$$

【注】重要结论：
$$\begin{cases}①(X_1,X_2)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)\Rightarrow X_1\sim N(\mu_1,\sigma_1^2)\text{，}X_2\sim N(\mu_2,\sigma_2{}^2)\\②X_1\sim N(\mu_1,\sigma_1^2)\text{，}X_2\sim N(\mu_2,\sigma_2^2)\text{，且}X_1X_2\text{独立}\Rightarrow(X_1,X_2)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,0)\\③(X_1,X_2)\sim N\Rightarrow k_1X_1+k_2X_2\sim N\ (k_1k_2\text{不全为}0)\\④(X_1,X_2)\sim N\Rightarrow(a_1X_1+a_2X_2,b_1Y_1+b_2Y_2)\sim N\text{，}\left(\begin{vmatrix}a_1&a_2\\b_1&b_2\end{vmatrix}\neq0\right)\end{cases}$$
，通过② $\Rightarrow$ ③，也即上面的(A).

6、设随机变量$X\sim t(n)$，则$Y=\dfrac{1}{X^2}$服从____.

【分析】具体化：$X_1\sim N(0,1)$，$Y_i\sim N(0,1)\Rightarrow Y_1^2+Y_2^2+\cdots+Y_n^2\sim\chi^2(n)\Rightarrow \dfrac{X_1}{\sqrt{\frac{Y_1^2+Y_2^2+\cdots+Y_n^2}{n}}}\sim t(n)$，$\Rightarrow X=\dfrac{X_1}{\sqrt{\frac{Y_1^2+Y_2^2+\cdots+Y_n^2}{n}}}\sim t(n)$

$\left(\text{题目中的}Y\text{就是个记号不用管他写成}Z=\dfrac{1}{X^2}\text{都行}\right)\Rightarrow \dfrac{1}{X^2}=\dfrac{1}{\left(\dfrac{X_1}{\sqrt{\frac{Y_1^2+Y_2^2+\cdots+Y_n^2}{n}}}\right)^2}=\dfrac{\frac{Y_1^2+Y_2^2+\cdots+Y_n^2}{n}}{X_1^2}=F(n,1).$

7、X_1，$X_2\cdots X_n$来自总体$N(0,1)$，$\bar X$为样本均值，S为样本方差，则( )

$(A)n\bar X\sim N(0,1)$　　$(B)nS^2\sim\chi^2(n)$　　$(C)\dfrac{(n-1)\bar X}{S}\sim t(n-1)$　　$(D)\dfrac{(n-1)X_1^2}{\sum_{i=2}^{n}X_i^2}\sim F(1,n-1)$

【分析】实际上对于这种题，先把抽样分布的结论写一下：就是上面的【总结一下】的内容，$X_i\sim N(\mu,\sigma^2)$，$\bar X\sim N\left(\mu,\dfrac{\sigma^2}{n}\right)$

$\xRightarrow{\text{对}X_i}\sum_{i=1}^{n}\left(\dfrac{X_i-\mu}{\sigma}\right)\sim\chi^2(n)$，$\sum_{i=1}^{n}\left(\dfrac{X_i-\bar X}{\sigma}\right)\sim\chi^2(n-1)=\dfrac{(n-1)S^2}{\sigma^2}$，$\xRightarrow{\text{对}\bar X}\dfrac{\bar X-\mu}{\frac{\sigma}{\sqrt n}}\sim N(0,1)$，$\dfrac{\bar X-\mu}{\frac{S}{\sqrt n}}\sim t(n-1)$，$\dfrac{\left(\bar X-\mu\right)^2}{\frac{S^2}{n}}\sim F(1,n-1)$

【观察】根据题目给的$X_i\sim N(0,1)$，带入上式，令$\mu=0$，$\sigma=1$

$\Rightarrow$ 看A，找带有N和$\bar X$的式子，$\dfrac{\bar X-0}{\frac{1}{\sqrt n}}\sim N(0,1)\Rightarrow\sqrt n\,\bar X\sim N(0,1)$，$A$错

看B，找带有$\chi^2(n)$和S^2的式子，发现，有一个$\dfrac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1)\xRightarrow{\text{题目给的}\sigma=1}(n-1)S^2\sim\chi^2(n-1)$，请问能否做变量替换$\Rightarrow nS^2\sim\chi^2(n)$

答，不可以，因为你的S和n是有关系的，当你变量替换令$n-1=n$时，S也发生了变化，所以$nS^2\sim\chi^2(n)$是错的，且注意S^2样本方差只和$n-1$在一起

看C，找有$\bar X$，S和t的，$\dfrac{\bar X-\mu}{\frac{S}{\sqrt n}}\sim t(n-1)$，带入$\mu=0\Rightarrow\dfrac{\bar X}{\frac{S}{\sqrt n}}\sim t(n-1)\Rightarrow\dfrac{\sqrt n\,\bar X}{S}\sim t(n-1)$，$C$错。

看D，这个显然自己构造，上面一个$X_1^2\sim\chi^2(1)$，下面$\sum_{i=2}^{n}X_i^2\sim\chi^2(n-1)\Rightarrow\dfrac{X_1^2/1}{\sum_{i=2}^{n}X_i^2/(n-1)}\sim F(1,n-1)$，即$\dfrac{(n-1)X_1^2}{\sum_{i=2}^{n}X_i^2}\sim F(1,n-1)D$对.

8、X_1，X_2，$\cdots X_n$来自正太总体$N(\mu,\sigma^2)$，$\bar X$是样本均值，S^2是样本方差，且有

$S_1^2=\dfrac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\bar X\right)^2$，$S_2^2=\dfrac{1}{n}\sum_{i=1}^{n}\left(X_i-\bar X\right)^2$，$S_3^2=\dfrac{1}{n-1}\sum_{i=1}^{n}(X_i-\mu)^2$，$S_4^2=\dfrac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2$则服从$t(n-1)$的分布是( )

$(A)t=\dfrac{\bar X-\mu}{S_1/\sqrt{n-1}}$　$(B)t=\dfrac{\bar X-\mu}{S_2/\sqrt{n-1}}$　$(C)t=\dfrac{\bar X-\mu}{S_3/\sqrt n}$　$(D)t=\dfrac{\bar X-\mu}{S_4/\sqrt n}$

【分析】想结论：$\dfrac{X_i-\mu}{\sigma}\sim N(0,1)$，$\sum_{i=1}^{n}\left(\dfrac{X_i-\mu}{\sigma}\right)^2\sim\chi^2(n)$，$\sum_{i=1}^{n}\left(\dfrac{X_i-\bar X}{\sigma}\right)^2\sim\chi^2(n-1)$，$\dfrac{\bar X-\mu}{\frac{\sigma}{\sqrt n}}\sim N(0,1)$，$\dfrac{\bar X-\mu}{\frac{S}{\sqrt n}}\sim t(n-1)$

产生t分布的就$\dfrac{\bar X-\mu}{\frac{S}{\sqrt n}}\sim t(n-1)$这一个，其中$S^2$是样本方差$\dfrac{1}{n-1}\sum_{i=1}^{n}\left(X-\bar X\right)^2$，所以上面 减去$\mu$的都不对，和样本方差不沾边，只有$AB$沾边

如果选A的话，A中的S_1就是咱这里的S，那么应该是$t=\dfrac{\bar X-\mu}{\frac{S_1}{\sqrt n}}\sim t(n-1)$，而他这里是$\dfrac{S_1}{\sqrt{n-1}}$所以$A$不对，选$B$.

分析B：B不是样本方差，但是用B给的这个式子可以去自己造分布，如根据$S_2\Rightarrow nS_2^2=\sum_{i=1}^{n}\left(X_i-\bar X\right)^2\xRightarrow{\text{考虑到}}\sum_{i=1}^{n}\left(\dfrac{X_i-\bar X}{\sigma}\right)^2\sim\chi^2(n-1).$

$\Rightarrow\dfrac{nS_2^2}{\sigma^2}=\sum_{i=1}^{n}\left(\dfrac{X_i-\bar X}{\sigma}\right)^2\sim\chi^2(n-1)\xRightarrow[\bar X\sim N\left(\mu,\frac{\sigma^2}{n}\right)]{\text{造}t\text{分布}}\dfrac{\frac{\bar X-\mu}{\sigma/\sqrt n}}{\sqrt{\frac{nS_2^2}{(n-1)\sigma^2}}}=\dfrac{\bar X-\mu}{S_2/\sqrt{n-1}}\sim t(n-1).B$对.

9、设$X\sim t(n)$，$Y\sim F(1,n)$，给定$0<\alpha<0.5$，常数c满足$P\{X>c\}=\alpha$，则$P\{Y>c^2\}=$ ____.

【分析】抽象问题具体化容易发现规律：$X=\dfrac{X_1}{\sqrt{\dfrac{X_2^2+X_3^2+\cdots+X_{n+1}^2}{n}}}\sim t(n)$，$Y=\dfrac{X_1^2/1}{\dfrac{X_2^2+X_3^2+\cdots+X_{n+1}^2}{n}}\sim F(1,n)$，发现$X^2=Y$

因此$P\{Y>c^2\}\Leftrightarrow P\{X^2>c^2\}\Leftrightarrow P\{X<-c\text{或}X>c\}\Rightarrow$ 又$t(n)$是类正太分布的偶函数，所以$P\{X<-c\text{或}X>c\}=2\alpha$.

10、设总体X的概率密度为：$f(x)=\dfrac{1}{2}e^{-|x|}$，$X_i$是来自总体的简单随机样本，样本方差是$S^2$，则$ES^2=$ ____.

【分析】看【常用结论】的第二条，无论何种样本只要期望是μ，方差是$\sigma^2\Rightarrow E\overline{X}=\mu$，$ES^2=\sigma^2$，$D\overline{X}=\dfrac{\sigma^2}{n}$，$DS^2=\dfrac{2\sigma^4}{n-1}$务必记住

因此$ES^2=\sigma^2$，就是求方差，$DX=EX^2-(EX)^2$，$EX=\displaystyle\int_{-\infty}^{+\infty}x\frac{1}{2}e^{-|x|}=0$，$EX^2=\dfrac{1}{2}\displaystyle\int_{-\infty}^{+\infty}x^2e^{-|x|}dx=\int_0^{+\infty}x^2e^{-x}dx=\Gamma(3)=2\Gamma(2)=2$

$\Rightarrow\sigma^2=DX=0+2=2.$

11、X_i来自总体$X\sim N(\mu,\sigma^2)$，记统计量$Y=\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2$，求$EY$，$DY$

【分析】Y这个形式也太像咱们的结论了吧，$\dfrac{X_i-\mu}{\sigma}\sim N(0,1)\Rightarrow\displaystyle\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2\sim\chi^2(n)\Rightarrow\sum_{i=1}^{n}\left(\frac{X_i-\overline{X}}{\sigma}\right)\sim\chi^2(n-1)$

$$\Rightarrow\frac{nY}{\sigma^2}=\sum_{i=1}^{n}\left(\frac{X_i-\overline{X}}{\sigma}\right)^2\sim\chi^2(n-1)\Rightarrow E\left(\frac{nY}{\sigma^2}\right)=n-1=\frac{n}{\sigma^2}EY\Rightarrow EY=\frac{(n-1)\sigma^2}{n}$$

$$D\left(\frac{nY}{\sigma^2}\right)=2(n-1)=\frac{n^2}{\sigma^4}DY\Rightarrow DY=\frac{2\sigma^4(n-1)}{n^2}.$$

12、设X_i来自总体$N(\mu,\sigma^2)$，$\overline{X}=\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}X_i$，$S^2=\dfrac{1}{n-1}\displaystyle\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2$，$T=\overline{X}^2-\dfrac{1}{n}S^2$，

(1)证明T是μ^2的无偏估计，(2)$\mu=0$，$\sigma=1$时，求DT.

【分析】$ET=E\overline{X}^2-\dfrac{1}{n}ES^2$，其中$E\overline{X}^2=\left(E\overline{X}\right)^2+D\overline{X}$，又$E\overline{X}=\mu$，$D\overline{X}=\dfrac{\sigma^2}{n}\Rightarrow E\overline{X}^2=\mu^2+\dfrac{\sigma^2}{n}$，$ES^2=\sigma^2$

$$\Rightarrow ET=\mu^2+\frac{\sigma^2}{n}-\frac{1}{n}\sigma^2=\mu.$$

(2)注意【常用结论】第一条后面有X_i来自总体$\overline{X}$和S独立$\Rightarrow DT=D\overline{X}^2+\dfrac{1}{n^2}DS^2$，其中$DS^2=\left.\dfrac{2\sigma^4}{n-1}\right|_{\sigma=1}=\dfrac{2}{n-1}$

$D\overline{X}^2$怎么求，见到平方想到卡方，$\overline{X}\sim N\left(\mu,\dfrac{\sigma^2}{n}\right)\Rightarrow\dfrac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\sim N(0,1)\xlongequal[\sigma=1]{\text{带入}\mu=0}\sqrt{n}\,\overline{X}\sim N(0,1)\Rightarrow\left(\sqrt{n}\,\overline{X}\right)^2\sim\chi^2(1)$

$$\Rightarrow n\overline{X}^2\sim\chi^2(1)\Rightarrow D\left(n\overline{X}^2\right)=2\Rightarrow D\overline{X}^2=\frac{2}{n^2}\text{，综上：}DT=\frac{2}{n^2}+\frac{2}{n^2(n-1)}=\frac{2n}{n^2(n-1)}=\frac{2}{n(n-1)}.$$

13、设X_i来自总体$N(0,1)$，$\overline{X}$为样本均值，记$Y_i=X_i-\overline{X}$，求DY_i，求$cov(Y_1,Y_n)$，求$P\{Y_1+Y_n\leq 0\}$

【分析】(1)有平方才想卡方，没平方自己做，$D\left(X_i-\overline{X}\right)=D\left(X_i-\dfrac{1}{n}(X_1+\cdots+X_i+\cdots+X_n)\right)=D\left(\left(1-\dfrac{1}{n}\right)X_i+\dfrac{1}{n}\displaystyle\sum_{k\neq i}^{n}X_k\right)$

$$\xlongequal{\text{现在独立了}}\left(1-\frac{1}{n}\right)^2DX_i+\frac{1}{n^2}\sum_{k\neq i}^{n}DX_k=\left(1-\frac{1}{n}\right)^2+\frac{n-1}{n^2}=\frac{(n-1)^2+n-1}{n^2}=\frac{n^2-n}{n^2}=\frac{n-1}{n}.$$

(2)$cov(Y_1,Y_n)=cov\left(X_1-\overline{X},X_n-\overline{X}\right)=cov(X_1,X_n)-cov\left(X_1,\overline{X}\right)-cov\left(\overline{X},X_n\right)+D\overline{X}$

其中$cov(X_1,X_n)=0$，$cov\left(X_1,\overline{X}\right)=cov\left(X_1,\dfrac{X_1+X_2+\cdots+X_n}{n}\right)=\dfrac{1}{n}DX_1=\dfrac{1}{n}=cov\left(\overline{X},X_n\right)$，$D\overline{X}=\dfrac{\sigma^2}{n}=\dfrac{1}{n}$

$\Rightarrow cov(Y_1,Y_n)=-\dfrac{2}{n}+\dfrac{1}{n}=-\dfrac{1}{n}$. 【注】当然(1)也可以$D\left(X_i-\overline{X}\right)=DX_i+D\overline{X}-2cov\left(X_i,\overline{X}\right)=1+\dfrac{1}{n}-2\cdot\dfrac{1}{n}=1-\dfrac{1}{n}$.

(3)$P\{Y_1+Y_n\leq 0\}\Rightarrow P\left\{X_1-\overline{X}+X_n-\overline{X}\right\}$，研究一下：$X_1+X_n-2\overline{X}=X_1+X_n-\dfrac{2}{n}(X_1+\cdots+X_n)$

$$=\left(1-\frac{2}{n}\right)(X_1+X_n)-\frac{2}{n}\sum_{k=2}^{n-1}X_i\overset{\substack{X_i\text{独立且服从正太分布}\\ \text{则其线性组合也服从}}}{\sim}N\text{，其中}E(Y_1+Y_n)=\left(1-\frac{2}{n}\right)\cdot 2EX-\frac{2}{n}(n-2)EX=0\Rightarrow P\{Y_1+Y_n\leq 0\}=\frac{1}{2}.$$

14、X_i来自总体$X\sim N(\mu,\sigma^2)$，记统计量$Y=\frac{1}{n}\sum_{i=1}^{n}|X_i-\mu|$，求$EY$，$DY$

【分析】见到| |有意识的把里面的东西捏在一起，也是求分布函数的时候的常用的方法。

令$X_i-\mu=U$，则$U\sim N(0,\sigma^2)\Rightarrow EY=\frac{1}{n}\sum_{i=1}^{n}E|U|$，其中$E|U|=\int_{-\infty}^{+\infty}|u|\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{u^2}{2\sigma^2}}du=\frac{2}{\sqrt{2\pi}\sigma}\int_0^{+\infty}ue^{-\left(\frac{u}{\sqrt{2}\sigma}\right)^2}du$

$=\frac{2\sqrt{2}\sigma}{\sqrt{2\pi}\sigma}\cdot\sqrt{2}\sigma\cdot\int_0^{+\infty}\frac{u}{\sqrt{2}\sigma}e^{-\left(\frac{u}{\sqrt{2}\sigma}\right)^2}d\frac{u}{\sqrt{2}\sigma}=\frac{2\sqrt{2}\sigma}{\sqrt{\pi}}\cdot\frac{1}{2}\cdot\Gamma(1)=\frac{\sqrt{2\pi}\sigma}{\pi}.\Rightarrow EY=\frac{1}{n}\cdot n\cdot\frac{\sqrt{2\pi}\sigma}{\pi}=\frac{\sqrt{2\pi}\sigma}{\pi}.$

$DY=\frac{1}{n^2}\sum_{i=1}^{n}D|U|$，其中$D|U|=EU^2-(E|U|)^2$，其中$EU^2\xlongequal[EX^2=DX+(EX)^2]{\text{这就别伽马函数了}}DU+(EU)^2=\sigma^2+0=\sigma^2$

$\Rightarrow D|U|=\sigma^2-\frac{2\pi\sigma^2}{\pi^2}=\sigma^2\left(1-\frac{2}{\pi}\right).\Rightarrow DY=\frac{1}{n^2}\cdot n\cdot\sigma^2\left(1-\frac{2}{\pi}\right)=\frac{\sigma^2}{n}\left(1-\frac{2}{\pi}\right).$

15、设$X_1\cdots X_{10}$来自总体$X\sim N(0,\sigma^2)$，$Y^2=\frac{1}{9}\sum_{i=2}^{10}X_i^2$，则( )

$(A)X_1^2\sim\chi^2(1)$　$(B)Y^2\sim\chi^2(9)$　$(C)\frac{X_1}{Y}\sim t(9)$　$(D)\frac{X_1^2}{Y^2}\sim F(9,1)$

【分析】还是先把总结论背一下：$X_i\sim N(\mu,\sigma^2)\Rightarrow\frac{X_i-\mu}{\sigma}\sim N(0,1)\Rightarrow\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2\sim\chi^2(n)\Rightarrow\sum_{i=1}^{n}\left(\frac{X_i-\overline{X}}{\sigma}\right)^2\sim\chi^2(n-1)$

$\overline{X}\sim N\left(\mu,\frac{\sigma^2}{n}\right)\Rightarrow\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\sim N(0,1)\Rightarrow\frac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}}\sim t(n-1)\Rightarrow\frac{(\overline{X}-\mu)}{\frac{S^2}{n}}\sim F(1,n-1)$，现在令$\mu=0$

看(A)，找带χ的，$\frac{X_1^2}{\sigma^2}\sim\chi^2(1)$，实际上也没必要去上面找，自己稍微一推就出来了，$\frac{X_1-0}{\sigma}\sim N(0,1)\Rightarrow\left(\frac{X_1}{\sigma}\right)^2\sim\chi^2(1)$，$A$✗

看(B)，还是自己根据题目的条件构造，$X_i\sim N(0,\sigma^2)\Rightarrow\frac{X_i-0}{\sigma}\sim N(0,1)\Rightarrow\sum_{i=2}^{10}\left(\frac{X_i}{\sigma}\right)^2\sim t(9)$，而$\frac{1}{\sigma^2}\sum_{i=2}^{10}X_i^2=\frac{9Y^2}{\sigma^2}\sim t(9)B$✗

看(C)，自己造完和答案比$\frac{\frac{X_1-0}{\sigma}}{\sqrt{\frac{9Y^2}{\sigma^2}/9}}\sim t(9)$，即$\frac{X_1}{Y}\sim t(9)$，$C$✓

看(D)，$\frac{\left(\frac{X_1-0}{\sigma}\right)^2}{\sum_{i=2}^{10}\left(\frac{X_i-0}{\sigma}\right)^2/9}\sim t(1,9)$，即$\frac{\frac{X_1^2}{\sigma^2}}{\frac{Y^2}{\sigma^2}}=\frac{X_1^2}{Y^2}\sim F(1,9)$. D✗

16、X_iY_i分别来自总体XY，且XY独立，且都服从$N(0,\sigma^2)$，则( )

$(A)\overline{X}-\overline{Y}\sim N(0,\sigma^2)$　$(B)S_X^2+S_Y^2\sim\chi^2(2n-2)$　$(C)\frac{\overline{X}-\overline{Y}}{\sqrt{S_X^2+S_Y^2}}\sim t(2n-2)$　$(D)\frac{S_X^2}{S_Y^2}\sim F(n-1,n-1)$

【分析】看A：$\overline{X}\sim N\left(0,\frac{\sigma^2}{n}\right)$，$\overline{Y}\sim N\left(0,\frac{\sigma^2}{n}\right)\Rightarrow E\left(\overline{X}-\overline{Y}\right)=0$，$D\left(\overline{X}-\overline{Y}\right)=\frac{2\sigma^2}{n}\Rightarrow\overline{X}-\overline{Y}\sim N\left(0,\frac{2\sigma^2}{n}\right)$ A✗

看B：$S_X^2=\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2$，由于$\frac{X_i-\mu}{\sigma}\sim N(0,1)\Rightarrow\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2\sim\chi^2(n)\Rightarrow\sum_{i=1}^{n}\left(\frac{X_i-\overline{X}}{\sigma}\right)^2\sim\chi^2(n-1)$

而$S_X^2\neq\sum_{i=1}^{n}\left(\frac{X_i-\overline{X}}{\sigma}\right)^2$，故$B$错，而应该是，由于$\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2=(n-1)S_X^2\Rightarrow\sum_{i=1}^{n}\left(\frac{X_i-\overline{X}}{\sigma}\right)^2=\frac{(n-1)S_X^2}{\sigma^2}\sim\chi^2(n-1)$

进而：$\frac{(n-1)}{\sigma^2}[S_X^2+S_Y^2]\sim\chi^2(2n-2)$ 再看C：自己构造去：$\frac{\frac{\overline{X}-\overline{Y}-0}{\frac{\sqrt{2}\sigma}{\sqrt{n}}}}{\sqrt{\frac{\frac{(n-1)S_X^2}{\sigma^2}+\frac{(n-1)S_Y^2}{\sigma^2}}{2n-2}}}\sim t(2n-2)$，即$\frac{\overline{X}-\overline{Y}}{\frac{S_X^2+S_Y^2}{\sqrt{n}}}\sim t(2n-2)$，$C$错.

看D：$\frac{(n-1)S_X^2/n-1}{(n-1)S_Y^2/n-1}\sim F(n-1,n-1)$，即$\frac{S_X^2}{S_Y^2}\sim F(n-1,n-1)$，$D$对.

【注】你上面那样推了一下S_X^2满足的分布，实际上没必要推啊，不是有一个$\sum_{i=1}^{n}\left(\frac{X_i-\overline{X}}{\sigma}\right)^2\sim\chi^2(n-1)=\frac{(n-1)S^2}{\sigma^2}$

$\xrightarrow{\text{马上得到}}\frac{(n-1)S_X^2}{\sigma^2}\sim\chi^2(n-1)$，我上面蓝色的地方自己推了一遍又，记住结论，没必要推，当然推了说明你掌握的好。

参数估计和评价标准

一、点估计
- 矩估计
  - 估计一个参数：
    - 直接令$\overline{X}=EX$
    - 若期望算出来是0，再令$\frac{1}{n}\sum_{i=1}^{n}X_i^2=E(X^2)$
  - 估计两个参数：
    - 先令$\overline{X}=EX$
    - 再令$\frac{1}{n}\sum_{i=1}^{n}X_i^2=E(X^2)$
- 最大似然估计[1]：
  - ①出现x_i这样一个观测值的概率是概率密度[2]那么大或者集中概率那么大，把每个观测对应的概率都连乘起来，就是似然函数
  - ②似然函数对其中的参数求偏导，令其为0
    - 若能解出参数，则解出来的参数就是所求的最大似然估计
    - 若解不出参数
      - 似然函数是参数的增函数，则参数越大越好
      - 似然函数是参数的减函数，则参数越小越好

【注1】似然定理：求$g(\theta)$这个函数的估计，就等于先把$\hat{\theta}$求出来，再带入函数中，只要这个函数是单调函数，比如求$\widehat{x^3}=\hat{x}^3$，见【预测】(2).

【注2】给你$F(x;\alpha,\beta)=\begin{cases}1-\left(\dfrac{\alpha}{x}\right)^{\beta}, & x>\alpha\\ 0, & x\leqslant\alpha\end{cases}$，$\alpha>0,\beta>1$求$\begin{cases}①\alpha=1\text{时，}\beta\text{的矩估计}\\ ②\alpha=1\text{时，}\beta\text{的最大似然估计}\\ ③\beta=2\text{时，}\alpha\text{的最大似然估计}\end{cases}$

【分析】让你求矩估计或者最大似然估计，<u>首先得把概率密度求出来</u>，因为他给的是分布函数. $f(x)=F'(x)=\begin{cases}\beta\alpha^{\beta}x^{-(\beta+1)}, & x>\alpha\\ 0, & x\leqslant\alpha\end{cases}$

①$\alpha=1$，$f(x)=\beta x^{-(\beta+1)}$，求矩估计就是$\overline{X}\overset{令}{=}EX$，求$EX=\int_{\alpha=1}^{+\infty}x\beta x^{-(\beta+1)}dx=\beta\int_{1}^{+\infty}x^{-\beta}dx=\left.\dfrac{\beta x^{-\beta+1}}{-\beta+1}\right|_{1}^{+\infty}=-\dfrac{\beta}{1-\beta}$

$\Rightarrow$ 令$\overline{X}=\dfrac{\beta}{\beta-1}\Rightarrow\hat{\beta}=\dfrac{\overline{X}}{\overline{X}-1}$.

②$L=\prod_{i=1}^{n}f(x_i)=\beta^n(x_1\cdots x_n)^{-(\beta+1)}$，$x_i>1$，$\ln L=n\ln\beta-(\beta+1)\sum_{i=1}^{n}\ln x_i\Rightarrow\dfrac{\partial\ln L}{\partial\beta}=\dfrac{n}{\beta}-\sum_{i=1}^{n}\ln x_i=0\Rightarrow\hat{\beta}=\dfrac{n}{\sum_{i=1}^{n}\ln x_i}$

③$L=\prod_{i=1}^{n}(2\alpha^2x_i^{-3})=2^n a^{2n}(x_1\cdots x_n)^{-3}$，$x_i>\alpha\Rightarrow\ln L=n\ln2+2n\ln a-3\sum_{i=1}^{n}\ln x_i\Rightarrow\dfrac{\partial\ln L}{\partial\alpha}=\dfrac{2n}{\alpha}>0$，

故L是关于α的增函数，所以α越大，L越大，又$x_1x_2\cdots x_n>\alpha$即$\alpha<x_i\Rightarrow\alpha<\min\{x_i\}$，取$\hat{\alpha}=\min\{x_i\}=x_{(1)}$

【2018】X_i来自总体$X\sim f(x,\sigma)=\dfrac{1}{2\sigma}e^{-\frac{|x|}{\sigma}}$，$\sigma>0$，$\hat{\sigma}$的最大似然估计量为$\hat{\sigma}$，求$\hat{\sigma}$，求$E\hat{\sigma}$，求$D\hat{\sigma}$

【分析】①$L=\prod_{i=1}^{n}\dfrac{1}{2\sigma}e^{-\frac{|x_i|}{\sigma}}=\dfrac{1}{2^n}\cdot\dfrac{1}{\sigma^n}e^{-\frac{|x_1|+|x_2|+\cdots+|x_n|}{\sigma}}\Rightarrow\ln L=-n\ln2-n\ln\sigma-\dfrac{\sum_{i=1}^{n}|x_i|}{\sigma}\Rightarrow\dfrac{\partial\ln L}{\partial\sigma}=-\dfrac{n}{\sigma}+\sum_{i=1}^{n}|x_i|\cdot\sigma^{-2}$

$=\dfrac{\sum_{i=1}^{n}|x_i|-n\sigma}{\sigma^2}=0\Rightarrow\hat{\sigma}=\dfrac{\sum_{i=1}^{n}|x_i|}{n}=\overline{|X|}$，②$E\hat{\sigma}=E\left(\dfrac{\sum_{i=1}^{n}|x_i|}{n}\right)=\dfrac{1}{n}\sum_{i=1}^{n}E|x_i|=\dfrac{1}{n}\cdot n\cdot E|x|=E|x|=\int_{-\infty}^{+\infty}|x|\dfrac{1}{2\sigma}e^{-\frac{|x|}{\sigma}}dx$

$=\dfrac{1}{2\sigma}\cdot2\cdot\int_{0}^{+\infty}xe^{-\frac{x}{\sigma}}dx=\sigma\int_{0}^{+\infty}\dfrac{x}{\sigma}e^{-\frac{x}{\sigma}}d\dfrac{x}{\sigma}=\sigma\Gamma(2)=\sigma$. ③$D\hat{\sigma}=D\left(\dfrac{\sum_{i=1}^{n}|x_i|}{n}\right)=\dfrac{1}{n^2}\sum_{i=1}^{n}D|x_i|=\dfrac{1}{n^2}\cdot n\cdot D|x|$

又$D|x|=Ex^2-(E|x|)^2$，其中$Ex^2=\int_{-\infty}^{+\infty}x^2\dfrac{1}{2\sigma}e^{-\frac{|x|}{\sigma}}dx-\sigma^2=\dfrac{2}{2\sigma}\int_{0}^{+\infty}x^2e^{-\frac{x}{\sigma}}dx=\dfrac{\sigma^3}{\sigma}\int_{0}^{+\infty}\left(\dfrac{x}{\sigma}\right)^2e^{-\frac{x}{\sigma}}d\dfrac{x}{\sigma}=\sigma^2\Gamma(3)=2\sigma^2$

$\Rightarrow D|x|=2\sigma^2-(E|x|)^2=2\sigma^2-\sigma^2=\sigma^2\Rightarrow D\hat{\sigma}=\dfrac{1}{n^2}\cdot n\cdot D|x|=\dfrac{\sigma^2}{n}$.

二、评价标准 $\begin{cases}\text{无偏性：狗估计的期望等于狗本身，即}E\,\widehat{狗}=狗\text{，叫做}\widehat{狗}\text{是狗的无偏估计}\left(E\hat{\theta}=\theta\text{，称}\hat{\theta}\text{是}\theta\text{的无偏估计}\right)\text{，对你求出来的估计求期望}\\ \text{有效性：如果}E\hat{\theta}_1=\theta\text{，}E\hat{\theta}_2=\theta\text{，但是}D\hat{\theta}_1<D\hat{\theta}_2\text{，称}\hat{\theta}_1\text{比}\hat{\theta}_2\text{更有效}\\ \text{一致相合性：如果}\hat{\theta}\text{依概率收敛到}\theta\text{，叫}\hat{\theta}\text{是}\theta\text{的一致相合估计，即}\lim\limits_{n\to\infty}P\left\{\left|\hat{\theta}-\theta\right|<\varepsilon\right\}=1\end{cases}$

【注】关于一致性：用两种方法说明 $\begin{cases}\text{辛钦大数定律：见下面第2题的(3)}\\ \text{切比雪夫不等式}\begin{cases}P\{|X-EX|\geqslant\varepsilon\}\leqslant\dfrac{DX}{\varepsilon^2}\to 0 \quad \Rightarrow \text{说明}\lim\limits_{n\to\infty}P=0\\ P\{|X-EX|<\varepsilon\}\geqslant 1-\dfrac{DX}{\varepsilon^2}\Rightarrow\text{说明}\lim\limits_{n\to\infty}P=1\end{cases}\end{cases}$

【评价标准补充练习】

1、设X_i来自总体$X\sim B(n,p)$，若$\overline{X}+kS^2$为np^2的无偏估计，则$k=$ ____.

【分析】狗是猫的无偏估计 $\Rightarrow$ 狗的期望等于猫，$E\left(\overline{X}+kS^2\right)=E\overline{X}+kES^2 \xlongequal[\overline{X}\sim F\left(\mu,\frac{\sigma^2}{n}\right)]{X\sim F\left(\mu,\sigma^2\right)} np+k\sigma^2=np+knpq=np+knp(1-p)\xlongequal{令}np^2$

$\Rightarrow k=\dfrac{np^2-np}{np(1-p)}=\dfrac{p-1}{1-p}=-1$.

2、X_i来自$X\sim f(x,\theta)=\begin{cases}\dfrac{1}{\theta}e^{-\frac{x}{\theta}},x>0\\ 0,\text{其他}\end{cases}$，记统计量$Z_1=\overline{X}$，$Z_2=n\min\{X_1\cdots X_n\}$

(1)说明Z_1是否是Z_2的无偏估计，

(2)若是无偏估计，比较Z_1Z_2的有效性

(3)Z_1是否是θ的一致估计

【分析】(1)就是证$EZ_1=Z_2$，$EZ_1=E\overline{X}=EX=\displaystyle\int_0^{+\infty}x\frac{1}{\theta}e^{-\frac{x}{\theta}}dx=\theta\int_0^{+\infty}\frac{x}{\theta}e^{-\frac{x}{\theta}}d\frac{x}{\theta}=\theta\Gamma(2)=\theta$

$Z_2=n\min\{X_1\cdots X_n\}$，$EZ_2=nE\min\{X_1\cdots X_n\}$，记$U=\min\{X_1\cdots X_n\}$，则分布函数$F_U(x)\xlongequal[\text{见〖多维随机变量〗}]{\text{回忆最小值}}1-[1-F(x)]^n$

其中$f(x)\xlongequal[\text{不要积分，立即得出}F(x)]{\text{（这是一个指数分布）}}\begin{cases}\lambda e^{-\lambda x}\text{，}x>0\\ 0\text{，其他}\end{cases}\Rightarrow F(x)=\begin{cases}1-e^{-\lambda x}\text{，}x>0\\ 0\text{，其他}\end{cases}\xRightarrow[\text{即}]{\text{对于此题}}F(x)=\begin{cases}1-e^{-\frac{x}{\theta}}\text{，}x>0\\ 0\text{，其他}\end{cases}$

$F_U(x)=1-[1-F(x)]^n=1-e^{-\frac{nx}{\theta}}$，$x>0\Rightarrow F_U(x)=\begin{cases}1-e^{-\frac{nx}{\theta}}\text{，}x>0\\ 0\text{，其他}\end{cases}\Rightarrow EZ_2=n\dfrac{\theta}{n}=\theta$

(2)$DZ_1=\dfrac{\sigma^2}{n}=\dfrac{1}{n}DX\xlongequal{\text{还是利用 指数分布}}\dfrac{1}{n}\cdot\theta^2=\dfrac{\theta^2}{n}$，$DZ_2=n^2D\min\{X_1\cdots X_n\}\xlongequal{\min\{Xi\}\text{也是指数分布}}n^2\cdot\dfrac{1}{\frac{n^2}{\theta^2}}=\theta^2$ 所以Z_1更有效.

(3)如果证明$\lim\limits_{n\to\infty}P\{|Z_1-\theta|<\varepsilon\}=1$，$P\{\theta-\varepsilon<Z_1<\varepsilon+\theta\}$就比较麻烦，所以证明依概率收敛就是用大数定律

$Z_1=\overline{X}=\dfrac{X_1+X_2+\cdots+X_n}{n}$，$X_i$独立同分布，期望存在，用辛钦大数定律，均值依概率收敛到均值的期望

$Z_1\xrightarrow{p}E\left(\dfrac{X_1+X_2+\cdots+X_n}{n}\right)=EX\xlongequal{X\text{是指数分布}}\theta$，证毕.

【注1】一般根据概率密度反求分布函数，你都先想一下，是不是常见的分布，直接得到分布函数，而不是去积分回去。

【注2】判断一致性，要考虑大数定律。

3、设总体X服从$[0,\theta]$上的均匀分布，θ未知$(\theta>0)$，$X_1X_2X_3$是取自X的样本

(1)$\hat{\theta}_1$是$k_1\max\{x_i\}$的无偏估计，求k_1值，并证明$\hat{\theta}_2$是$4\min\{x_i\}$的无偏估计，

(3)比较$\hat{\theta}_1$和$\hat{\theta}_2$的有效性.

【分析】关于最值的问题一定要涉及到分布函数来求，先写出$X\sim f(x)=\begin{cases}\dfrac{1}{\theta}, & 0\leqslant x\leqslant\theta\\ 0, & \text{其他}\end{cases}$，$X\sim F(x)=\begin{cases}0, & x<0\\ \dfrac{x}{\theta}, & 0\leqslant x<\theta\\ 1, & x\geqslant\theta\end{cases}$

令$Y=\max\{x_i\},Z=\min\{x_i\}\Rightarrow F_Y(x)=[F(x)]^3\Rightarrow f_Y(x)=3F^2(x)f(x)=\begin{cases}3\left(\dfrac{x}{\theta}\right)^2\dfrac{1}{\theta}, & 0\leqslant x<\theta\\ 0, & \text{其他}\end{cases}$

$F_Z(x)=1-[1-F(x)]^3\Rightarrow f_Z(x)=3[1-F(x)]^2f(x)=\begin{cases}3\left(1-\dfrac{x}{\theta}\right)^2\dfrac{1}{\theta}, & 0\leqslant x<\theta\\ 0, & \text{其他}\end{cases}$

由于$\hat{\theta}_1$是$k_1\max\{x_i\}$的无偏估计$\Leftrightarrow\hat{\theta}_1=k_1EY=k_1\displaystyle\int_0^\theta x\cdot3\left(\frac{x}{\theta}\right)^2\frac{1}{\theta}dx=\frac{3}{4}k_1\theta=\hat{\theta}_1\Rightarrow k_1=\frac{4}{3}$.

$\hat{\theta}_2=E(4\min\{x_i\})=4EZ=4\displaystyle\int_0^\theta x3\left(1-\frac{x}{\theta}\right)^2\frac{1}{\theta}dx=\theta$，$\left(EZ=\dfrac{1}{4}\theta\right)$证毕.

(2)$D\hat{\theta}_1=D\left(\dfrac{4}{3}Y\right)=\dfrac{16}{9}DY$，其中$DY=EY^2-(EY)^2=\displaystyle\int_0^\theta x^2\cdot3\left(\frac{x}{\theta}\right)^2\frac{1}{\theta}dx-\left(\frac{3}{4}\theta\right)^2=\frac{3}{5}\theta^2-\frac{9}{16}\theta^2=\frac{3}{80}\theta^2$

$\Rightarrow D\hat{\theta}_1=\dfrac{16}{9}\cdot\dfrac{3}{80}\theta^2=\dfrac{\theta^2}{15}$.

$D\hat{\theta}_2=D(4Z)=16DZ=16(EZ^2-(EZ)^2)=16\cdot\dfrac{3}{80}\theta^2=\dfrac{3}{5}\theta^2>\dfrac{\theta^2}{15}\Rightarrow\hat{\theta}_1$更有效.

4、设$X_1X_2\cdots X_n$来自总体X，$EX=\mu$，$DX=\sigma^2$，证明统计量$Y=\dfrac{2}{n(n+1)}\displaystyle\sum_{i=1}^n iX_i$是$\mu$的无偏相合估计量.

【分析】先说明无偏性：$EY=E\left(\dfrac{2}{n(n+1)}\displaystyle\sum_{i=1}^n iX_i\right)=\frac{2}{n(n+1)}\sum_{i=1}^n iEX_i=\frac{2\mu}{n(n+1)}\sum_{i=1}^n i=\frac{2\mu}{n(n+1)}\cdot\frac{n(n+1)}{2}=\mu$，$EY=\mu\therefore Y$是$\mu$的无偏估计

证明一致相合性：考虑$\begin{cases}\text{辛钦大数定律}\\ \text{切比雪夫不等式}\end{cases}$，由切比雪夫不等式$P\{|Y-EY|\geqslant\varepsilon\}\leqslant\dfrac{DY}{\varepsilon^2}$，

其中$DY=D\left(\dfrac{2}{n(n+1)}\displaystyle\sum_{i=1}^n iX_i\right)=\frac{4}{n^2(n+1)^2}\sum_{i=1}^n i^2DX_i=\frac{4\sigma^2}{n^2(n+1)^2}\sum_{i=1}^n i^2=\frac{4\sigma^2}{n^2(n+1)^2}\cdot\frac{n(n+1)(2n+1)}{6}=\frac{2n(2n+1)}{3n^2(n+1)}\xrightarrow{n\to\infty}0$

即$\lim\limits_{n\to\infty}P\{|Y-EY|\geqslant\varepsilon\}=0$，而$EY=\mu$，故$Y$是$\mu$的一致相合估计.

【注】上面第2题的(3)，是证明$Z_1=\dfrac{X_1+X_2+\cdots+X_n}{n}$是$\theta$的无偏估计，就是证$\lim\limits_{n\to\infty}P\{|Z_1-\theta|<\varepsilon\}=1$，要用辛钦大数定律是均值依概率收敛到均值的期望，也就是你得保证Z_1是样本均值，而θ正好是样本均值的期望，于是证明了$Z_1\to\theta$，而这里的Z_1正好就是X_i的样本均值，因此可以由辛钦大数定律，直接得到$Z_1\to EZ_1$，而$EZ_1=E\left(\dfrac{X_1+X_2+\cdots+X_n}{n}\right)=\dfrac{1}{n}\cdot n\cdot EX=\theta$. 正好是$|Z_1-\theta|$的$\theta$于是证毕. 这个第4题证明一致性，用辛钦大数定律的话，就是证明$P\{|Y-\mu|<\varepsilon\}=1$，如果说明$Y$是样本均值，$\mu$是样本均值的期望，那就用辛钦大数定律证出来了，但$Y$是样本均值不是很好看吧，所以考虑第二个方法用切比雪夫不等式，因为切比雪夫不等式减去的就是他自己期望，不用管他是不是样本均值，所以用$P\{|Y-EY|\geqslant\varepsilon\}\leqslant\dfrac{DX}{\varepsilon^2}\Rightarrow\lim\limits_{n\to\infty}P=0$，就证出来了。或者用$P\{|Y-EY|\leqslant\varepsilon\}\geqslant1-\dfrac{DX}{\varepsilon^2}\to1$

【点估计补充练习】

1、设总体$X \sim f(x;a)=\begin{cases}\dfrac{4x^2}{a^3\sqrt{\pi}}e^{-\frac{x^2}{a^2}}, & x>0\\ 0 & , x\leqslant 0\end{cases}$ $(a>0)$，x_i是从X中取出的观测值，求a的矩估计值。

【分析】$\overline{X}\xlongequal{令}EX=\int_{-\infty}^{+\infty}xf(x)dx=\dfrac{2a}{\sqrt{\pi}} \Rightarrow \hat{a}=\dfrac{\sqrt{\pi}}{2}\overline{x}$

【注】求矩估计值小写x，求矩估计量大写X.

2、设总体X的概率密度$f(x)=\dfrac{1}{2\theta}e^{-\frac{|x|}{\theta}},\theta>0$，$X_i$是取自$X$的样本，求$\theta$的矩估计量.

【分析】$\overline{X}\xlongequal{令}EX=\int_{-\infty}^{+\infty}xf(x)dx\xlongequal{奇函数}0$（失效）$\Rightarrow \dfrac{1}{n}\sum\limits_{i=1}^{n}X_i^2\xlongequal{令}EX^2=\int_{-\infty}^{+\infty}x^2f(x)dx=2\theta^2 \Rightarrow \hat{\theta}=\sqrt{\dfrac{1}{2n}\sum\limits_{i=1}^{n}X_i^2}$.

3、两个参数：总体$X\sim f(x;\theta,\lambda)=\begin{cases}\dfrac{1}{\lambda}e^{-\frac{1}{\lambda}(x-\theta)}, & x\geqslant\theta\\ 0，其他\end{cases}$ $(\lambda>0)$，X_i来自X，求$\lambda\theta$的矩估计量.

【分析】$\overline{X}\xlongequal{令}EX=\int_{-\infty}^{+\infty}xf(x)dx=\lambda+\theta$；$\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i^2\xlongequal{令}EX^2=\int_{-\infty}^{+\infty}x^2f(x)dx=\lambda^2+(\lambda+\theta)^2$

$$\Rightarrow \begin{cases}\overline{X}=\lambda+\theta\\ \dfrac{1}{n}\sum\limits_{i=1}^{n}X_i^2=\lambda^2+(\lambda+\theta)^2\end{cases} \Rightarrow \begin{cases}\hat{\lambda}=\sqrt{\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i^2-\overline{X}}\\ \hat{\theta}=\overline{X}-\hat{\lambda}\end{cases}.$$

4、最大似然估计预测题：设X_i来自总体$X\sim B(m,p)$，m已知，p未知，求X的期望μ的最大似然估计.

【分析】由于$X\sim B(m,p) \Rightarrow EX=mp$，故$\hat{\mu}=m\hat{p}$，下面求$\hat{p}$

由于$P\{X=k\}=C_m^k p^k q^{m-k} \Rightarrow P\{X=x_i\}=C_m^{x_i}p^{x_i}(1-p)^{m-x_i} \Rightarrow$ 似然函数：$L=\prod\limits_{i=1}^{n}\left[C_m^{x_i}p^{x_i}(1-p)^{m-x_i}\right]$

$$\Rightarrow \ln L=\sum_{i=1}^{n}\ln\left(C_m^{x_i}p^{x_i}(1-p)^{m-x_i}\right)=\sum_{i=1}^{n}\left[\ln C_m^{x_i}+x_i\ln p+(m-x_i)\ln(1-p)\right]=\sum_{i=1}^{n}\ln C_m^{x_i}+\ln p\sum_{i=1}^{n}x_i+\ln(1-p)\sum_{i=1}^{n}(m-x_i)$$

$$=\sum_{i=1}^{n}\ln C_m^{x_i}+\ln p\sum_{i=1}^{n}x_i+\ln(1-p)\left[mn-\sum_{i=1}^{n}x_i\right]$$

$$\Rightarrow \frac{d\ln L}{dp}=\frac{\sum\limits_{i=1}^{n}x_i}{p}-\frac{mn-\sum\limits_{i=1}^{n}x_i}{1-p}\xlongequal{令}0 \Rightarrow \hat{p}=\frac{\sum\limits_{i=1}^{n}x_i}{mn}=\frac{\overline{X}}{m} \Rightarrow \hat{\mu}=m\hat{p}=\overline{X}.$$

【注】实际上，$L=\prod\limits_{i=1}^{n}\left[C_m^{x_i}p^{x_i}(1-p)^{m-x_i}\right]\xlongequal[当然也可以按照上面的方式]{不难直接写出}\prod\limits_{i=1}^{n}C_m^{x_i}\cdot p^{\sum\limits_{i=1}^{n}x_i}\cdot(1-p)^{mn-\sum\limits_{i=1}^{n}x_i}$

$$\Rightarrow \ln L=\ln\left(\prod_{i=1}^{n}C_m^{x_i}\right)+\ln p\cdot\sum_{i=1}^{n}x_i+\left(mn-\sum_{i=1}^{n}x_i\right)\ln(1-p)$$

5、最大似然估计预测题：设X_i来自总体$X\sim P(\lambda)$，求X的期望μ的最大似然估计.

【分析】泊松分布$X\sim P\{X=k\}=\dfrac{\lambda^k}{k!}e^{-\lambda} \Rightarrow P\{X=x_i\}=\dfrac{\lambda^{x_i}}{x_i!}e^{-\lambda} \Rightarrow EX=\mu=\lambda$(泊松分布的期望方差都是$\lambda$)故$\hat{\mu}=\hat{\lambda}$，下面求$\hat{\lambda}$

$$L=\prod_{i=1}^{n}\left[\frac{\lambda^{x_i}}{k!}e^{-\lambda}\right]=\frac{1}{(k!)^n}\lambda^{\sum\limits_{i=1}^{n}x_i}\cdot e^{-\lambda n} \Rightarrow \ln L=\ln\left(\frac{1}{(k!)^n}\right)+\sum_{i=1}^{n}x_i\cdot\ln\lambda-\lambda n \Rightarrow \frac{d\ln L}{d\lambda}=\frac{\sum\limits_{i=1}^{n}x_i}{\lambda}-n\xlongequal{令}0$$

$$\Rightarrow \hat{\lambda}=\frac{\sum\limits_{i=1}^{n}x_i}{n}=\overline{X} \Rightarrow \hat{\mu}=\hat{\lambda}=\overline{X}.$$

6、总体$X\sim f(x)=\begin{cases}\frac{1}{\theta}e^{-\frac{x-c}{\theta}}, & x\geqslant c,\ \theta>0\\ 0, & 其他\end{cases}$，其中$\theta c$是未知参数，抽取$n$个排序：$x_1\leqslant x_2\leqslant\cdots\leqslant x_n$，求$\theta$和$c$的最大似然估计值.

【分析】由于$X\sim f(x)$长上面那样 $\Rightarrow$ $f(x_i)=\begin{cases}\frac{1}{\theta}e^{-\frac{x_i-c}{\theta}}, & x_i\geqslant c,\ \theta>0\\ 0, & 其他\end{cases}$，实际上只写$f(x_i)=\frac{1}{\theta}e^{-\frac{x_i-c}{\theta}}$即可

$$\Rightarrow \text{似然函数：} L=\prod_{i=1}^{n}f(x_i)=\prod_{i=1}^{n}\left[\frac{1}{\theta}e^{-\frac{x_i-c}{\theta}}\right]=\frac{1}{\theta^n}\cdot e^{-\frac{\sum_{i=1}^{n}x_i-nc}{\theta}}\Rightarrow \ln L=-n\ln\theta-\frac{\sum_{i=1}^{n}x_i-nc}{\theta}$$

$$\Rightarrow \frac{\partial\ln L}{\partial\theta}=-\frac{n}{\theta}+\frac{\sum_{i=1}^{n}x_i-nc}{\theta^2}\overset{令}{=\!=}0\Rightarrow\hat{\theta}=\frac{\sum_{i=1}^{n}x_i-nc}{n}=\overline{x}-c.$$

$\Rightarrow \frac{\partial\ln L}{\partial c}=\frac{n}{\theta}>0\Rightarrow L$关于$c$是单增函数，所以$c$越大$L$越大，越最大似然，越好，而$x_i\geqslant c\Rightarrow c\leqslant x_i\Rightarrow c\leqslant\min\{x_i\}$取$\hat{c}=x_{(1)}$

综上：$\begin{cases}\hat{\theta}=\overline{x}-\hat{c}\\ c=x_{(1)}\end{cases}$.

【预测】设总体X服从对数正太分布，即$\ln X\sim N(\mu,\sigma^2)$，$X_i$来自$X$，(1)求$\mu$，$\sigma^2$的最大似然估计量，并讨论无偏性　(2)求$EX$的最大似然估计量

【分析】回忆最大似然估计怎么做，是取一个x_i，一个x_i取到的概率密度或者集中概率那么大，把他们连乘起来就算似然函数

所以现在你得知道我一个x_i，他取到的概率是多大，显然这是连续型不涉及集中概率的问题，因此应该首先把x的概率密度求出来。

$\ln X\sim N(\mu,\sigma^2)$这是$X$的函数服从的分布，所以要把$x$从其函数中摘出来，令$Y=\ln X\sim N(\mu,\sigma^2)$，则$X=e^Y$

$$则F_X(x)=P\{X\leqslant x\}=P\{e^Y\leqslant x\}=\begin{cases}0, & x\leqslant 0\\ P\{Y\leqslant\ln x\}, & x>0\end{cases}=\begin{cases}0, & x\leqslant 0\\ P\left\{\frac{Y-\mu}{\sigma}\leqslant\frac{\ln x-\mu}{\sigma}\right\}=\Phi\left(\frac{\ln x-\mu}{\sigma}\right), & x>0\end{cases}$$

$$\Rightarrow X\sim F_X(x)=\begin{cases}\Phi\left(\frac{\ln x-\mu}{\sigma}\right), & x>0\\ 0, & x\leqslant 0\end{cases}\Rightarrow X\sim f(x)=\begin{cases}\varphi\left(\frac{\ln x-\mu}{\sigma}\right)\cdot\frac{1}{\sigma x}, & x>0\\ 0, & x\leqslant 0\end{cases}\Rightarrow X\sim f(x)=\begin{cases}\frac{1}{\sqrt{2\pi}}e^{-\frac{\left(\frac{\ln x-\mu}{\sigma}\right)^2}{2}}\cdot\frac{1}{\sigma x}, & x>0\\ 0, & x\leqslant 0\end{cases}$$

$$即X\sim f(x)=\begin{cases}\frac{1}{\sqrt{2\pi}\sigma x}e^{-\frac{\left(\frac{\ln x-\mu}{\sigma}\right)^2}{2}}, & x>0\\ 0, & x\leqslant 0\end{cases}\Rightarrow f(x_i)=\begin{cases}\frac{1}{\sqrt{2\pi}\sigma x_i}e^{-\frac{\left(\frac{\ln x_i-\mu}{\sigma}\right)^2}{2}}, & x>0\\ 0, & x\leqslant 0\end{cases}$$

$$\Rightarrow \text{似然函数：} L=\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n\prod_{i=1}^{n}\frac{1}{x_i}\cdot e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(\ln x_i-\mu)^2}\Rightarrow \ln L=-n\ln\left(\sqrt{2\pi}\sigma\right)+\sum_{i=1}^{n}\ln\frac{1}{x_i}-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(\ln x_i-\mu)^2$$

后面要对σ^2这个整体求导，因此化简$\ln L=-n\ln\left(\sqrt{2\pi\sigma^2}\right)+\sum_{i=1}^{n}\ln\frac{1}{x_i}-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(\ln x_i-\mu)^2$

$$=-n\left(\ln\sqrt{2\pi}+\frac{1}{2}\ln\sigma^2\right)+\sum_{i=1}^{n}\ln\frac{1}{x_i}-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(\ln x_i-\mu)^2=-\frac{n}{2}\ln 2\pi-\frac{n}{2}\ln\sigma^2-\sum_{i=1}^{n}\ln x_i-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(\ln x_i-\mu)^2$$

$$\Rightarrow \frac{\partial\ln L}{\partial\mu}=\frac{1}{\sigma^2}\sum_{i=1}^{n}(\ln x_1-\mu)=\frac{1}{\sigma^2}\left[\sum_{i=1}^{n}\ln x_i-n\mu\right]\overset{令}{=\!=}0\Rightarrow \underline{\hat{\mu}=\frac{\sum_{i=1}^{n}\ln x_i}{n}};\ \frac{\partial\ln L}{\partial\sigma^2}=-\frac{n}{2}\cdot\frac{1}{\sigma^2}+\frac{1}{2}\cdot\frac{\sum_{i=1}^{n}(\ln x_i-\mu)^2}{\sigma^4}\overset{令}{=\!=}0$$

$$\Rightarrow \hat{\sigma}^2=\frac{\sum_{i=1}^{n}(\ln x_i-\mu)^2}{n}\Rightarrow \underline{\hat{\sigma}^2=\frac{\sum_{i=1}^{n}(\ln x_i-\hat{\mu})^2}{n}}.$$（最后把x_i写成X_i）

$$E\hat{\mu}=E\frac{1}{n}\sum_{i=1}^{n}\ln X_i=E\left(\frac{1}{n}\sum_{i=1}^{n}Y_i\right)=E\overline{Y}=\mu,\ \text{无偏；}\ E\hat{\sigma}^2=E\left(\frac{1}{n}\sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)^2\right)=E\left(\frac{n-1}{n}\cdot\frac{1}{n-1}\sum_{i=1}^{n}\left(Y_i-\overline{Y}\right)^2\right)=\frac{n-1}{n}ES_Y^2=\frac{n-1}{n}\sigma^2.$$

(2)欲求EX的估计，若有$EX=f(\mu,\sigma^2)$，则EX估计$=f\left(\hat{\mu},\hat{\sigma}^2\right)$，则$\hat{\mu}$和$\hat{\sigma}^2$已经求出，现求$EX=Ee^Y=\int_{-\infty}^{+\infty}e^y\frac{1}{\sqrt{2\pi}\sigma}\cdot e^{-\frac{(y-\mu)^2}{2\sigma^2}}dy=e^{\mu+2\sigma^2}$

$$\Rightarrow \widehat{EX}=e^{\hat{\mu}+2\hat{\sigma}^2}.\text{【注】}\int_{-\infty}^{+\infty}e^x\frac{1}{\sqrt{2\pi}\sigma}\cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty}e^{\frac{2\sigma^2x-(x-\mu)^2}{2\sigma^2}}dx=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty}e^{\frac{-\left[x-\left(\mu+\sigma^2\right)\right]^2+\left(\mu+\sigma^2\right)^2-\mu^2}{2\sigma^2}}dx$$

$$=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty}e^{-\frac{(x-a)^2}{2\sigma^2}}\cdot e^{\frac{2\mu\sigma^2+\sigma^4}{2\sigma^2}}dx=e^{\mu+\frac{\sigma^2}{2}}\cdot\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-数)^2}{2\sigma^2}}=e^{\mu+\frac{\sigma^2}{2}}\cdot 1=e^{\mu+\frac{\sigma^2}{2}}.$$

区间估计假设检验

【区间估计】 区间估计就是在σ^2已知或者未知的情况下，根据置信度去估计μ落在哪个区间。

估计的是μ，μ落在以$\overline{X}$为中心的双侧置信区间为$\left(\overline{X}-\triangle,\overline{X}+\triangle\right)\Rightarrow\begin{cases}若\sigma^2已知：\triangle=z_{\frac{\alpha}{2}}\cdot\dfrac{\sigma}{\sqrt{n}}\\ 若\sigma^2未知：\triangle=t_{\frac{\alpha}{2}}(n-1)\cdot\dfrac{S}{\sqrt{n}}\end{cases}$

【注1】若是单侧置信区间则把$\dfrac{\alpha}{2}$换成α即可

【注2】注意σ未知的时候，用$t_{\frac{\alpha}{2}}$估计的时候，多了一个$t_{\frac{\alpha}{2}}(n-1)$

【注3】$1-\alpha$叫置信水平也叫置信度，α叫显著性水平，也是犯第一类错误的概率.

【例1】设总体$X\sim N(\mu,\sigma^2)$，且σ^2已知，则在给定样本容量n以及置信度$1-\alpha$的情况下，未知参数μ的置信区间长度随着$\overline{X}$的增加而__不变__.

【分析】估计μ的置信区间 $\Rightarrow$ 以$\overline{X}$为中心，现在σ已知，μ落在$\left(\overline{X}-z_{\frac{\alpha}{2}}\dfrac{\sigma}{\sqrt{n}}，\overline{X}+z_{\frac{\alpha}{2}}\dfrac{\sigma}{\sqrt{n}}\right)$，置信区间长度为：$2z_{\frac{\alpha}{2}}\dfrac{\sigma}{\sqrt{n}}$，和$\overline{X}$没关系。

【例2】设总体$X\sim N(\mu,\sigma^2)$，且σ^2已知，n是样本容量，μ是未知参数，则μ的等尾双侧置信区间长度L与置信度$1-\alpha$的关系是，当$1-\alpha\downarrow$时L__$\downarrow$__.

【分析】和上个题一样，$L=2z_{\frac{\alpha}{2}}\dfrac{\sigma}{\sqrt{n}}$，$1-\alpha\downarrow\Rightarrow\alpha\uparrow$，$\alpha$增大，所以标准正太分布右侧的面积变大，则$z_{\frac{\alpha}{2}}\downarrow$，所以$L=2z_{\frac{\alpha}{2}}\dfrac{\sigma}{\sqrt{n}}\downarrow$

【例3】设总体$X\sim N(\mu,\sigma^2)$，其中μ，σ未知，先从X中取出16个样本，得到样本均值$\overline{x}=20$，样本标准差$s=1$，求μ的置信水平为0.9的置信区间.

【分析】现在σ不知道了，所以估计μ，就用$t_{\frac{\alpha}{2}}\dfrac{\sigma}{\sqrt{n}}$，于是，$\mu\in\left(\overline{x}-t_{\frac{\alpha}{2}}(n-1)\dfrac{S}{\sqrt{n}},\overline{x}+t_{\frac{\alpha}{2}}(n-1)\dfrac{S}{\sqrt{n}}\right)$，由于$1-\alpha=0.9\Rightarrow\alpha=0.1$

所以$\mu\in\left(20-t_{0.05}(15)\dfrac{1}{4}，20+t_{0.05}(15)\dfrac{1}{4}\right)$.

【假设检验】 区间估计是估计μ落在以$\overline{X}$为中心的哪个区间里面，假设检验是，我给你一个μ_0，问你是不是接受我给你的μ_0，

你接受或者拒绝的标准是什么？你就根据我给你的条件，算一下你的$\overline{X}$是不是落在，以我给你的μ_0为中心的哪个区间里面，没落在，就拒绝μ_0.

一、双边假设：

①σ^2已知，μ未知：如果你的$\overline{X}$落在$\left(\mu_0-z_{\frac{\alpha}{2}}\dfrac{\sigma}{\sqrt{n}}，\mu_0+z_{\frac{\alpha}{2}}\dfrac{\sigma}{\sqrt{n}}\right)$里面，那你就接受我给你的$\mu_0$，否则就拒绝。

②σ^2未知，μ未知：如果你的$\overline{X}$落在$\left(\mu_0-t_{\frac{\alpha}{2}}(n-1)\dfrac{S}{\sqrt{n}}，\mu_0+t_{\frac{\alpha}{2}}(n-1)\dfrac{S}{\sqrt{n}}\right)$里面，你就接受，落在外面就拒绝$\mu_0$。

【注】和区间估计很像，只不过区间估计是给你$\overline{X}$，让你估计$\mu\in\left(\overline{X}-\triangle,\overline{X}+\triangle\right)$，

可以把假设检验形式上理解为：现在是给你μ_0，让你估计$\overline{X}\in(\mu_0-\triangle,\mu_0+\triangle)$

二、单边假设：

①如果我给你一个$\mu\leqslant\mu_0$，那么，你可以把你接受的范围向右开拓一个$\triangle$

1.如果σ^2已知，μ未知，则如果你测出来的$\overline{X}\in\left(-\infty，\mu_0+z_\alpha\dfrac{\sigma}{\sqrt{n}}\right)$，那你就接受，否则就拒绝。注意单边写的是$z_\alpha$，把双边的$\dfrac{\alpha}{2}$换成$\alpha$

2.如果σ^2未知，μ未知，则如果你测出来的$\overline{X}\in\left(-\infty,\mu_0+t_\alpha(n-1)\dfrac{S}{\sqrt{n}}\right)$，那你就接受，否则就拒绝。

①如果我给你一个$\mu\geqslant\mu_0$，那么，你可以把你接受的范围向左开拓一个$\triangle$

1.如果σ^2已知，μ未知，则如果你测出来的$\overline{X}\in\left(\mu_0-z_\alpha\dfrac{\sigma}{\sqrt{n}}，+\infty\right)$，那你就接受，否则就拒绝。

2.如果σ^2未知，μ未知，则如果你测出来的$\overline{X}\in\left(\mu_0-t_\alpha(n-1)\dfrac{S}{\sqrt{n}}，+\infty,\right)$，那你就接受，否则就拒绝。

【假设检验练习】

【例】已知总体$X\sim N(\mu,\sigma^2)$，从中抽取16个样本，测得样本均值为$\overline{x}=10$，样本方差为$S^2=0.16$，$t_{0.025}(15)=2.132$.

(1)求总体均值μ的置信水平为0.95的置信区间 (2)在显著性水平为0.05下假设检验H_0：$\mu=9.7$，H_1：$\mu\neq 9.7$

【分析】(1)σ^2未知做区间估计：其中$1-\alpha=0.95\Rightarrow\alpha=0.05$，所以$\mu\in\left(\overline{X}-t_{\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}},\ \overline{X}+t_{\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}}\right)$

$\Rightarrow\ \mu\in\left(10-t_{0.025}(15)\frac{0.4}{4},10+t_{0.025}(15)\frac{0.4}{4}\right)$，即$\mu\in(9.7868,10.2132)$

(2)现在给了你$\mu_0=9.7$，所以只要$\overline{X}\in\left(\mu_0-t_{\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}},\ \mu_0+t_{\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}}\right)$，就接受我给你的9.7，否则就拒绝

$\overline{X}\in(9.4868,9.9132)$就接受否则拒绝，又$\overline{x}=10\notin(9.4868,9.9132)$，所以拒绝你的9.7，即拒绝$H_0$.

【两类错误】 $\begin{cases}\text{第一类错误：弃真的概率为}\alpha=P\{\text{但是拒绝了}H_0\mid\text{本应该接受}H_0\}\\ \text{第二类错误：取伪的概率}\beta=P\{\text{但是你接受了}H_0\mid\text{本应该拒绝}H_0\}\end{cases}$

【注】弃真的概率α就是区间估计里面的显著性水平α，由于弃真的后果很严重，所以一定要让弃真的概率尽可能的小，即让α尽量小，进而保护H_0

【例1】设总体$X\sim P(\lambda)$，X_1,X_2取自总体，假设检验问题H_0：$\lambda=0.5$，H_1：$\lambda=1$的拒绝域为$W=\{X_1+X_2\geqslant 2\}$

求该检验犯第一类错误和犯第二类错误的概率。

【分析】犯第一类错误就是你本来应该接受H_0，但是你却拒绝了他，拒绝了他就是落在拒绝域里面了，即$\alpha=P\{X_1+X_2\geqslant 2\mid H_0\}$

这个H_0就是λ取0.5，即$X\sim P(0.5)\Rightarrow X_1+X_2\sim P(1)\Rightarrow\alpha=P\{X_1+X_2\geqslant 2\mid H_0\}=1-P\{X_1+X_2<2|H_0\}$

$=1-(P\{X_1+X_2=1\mid H_0\}+P\{X_1+X_2=0\mid H_0\})=1-\left(\frac{\lambda^k}{k!}e^{-\lambda}\Big|_{\lambda=1,k=1}+\frac{\lambda^k}{k!}e^{-\lambda}\Big|_{\lambda=1,k=0}\right)=1-(e^{-1}+e^{-1})=1-\frac{2}{e}.$

犯第二类错误，就是你本应该拒绝H_0，但是你接受了$H_0\Rightarrow\beta=P\{X_1+X_2<2\mid H_1\}$，现在你帘子后面是$H_1$就导致了$X\sim P(1)$

$\Rightarrow\ X_1+X_2\sim P(2)\Rightarrow\beta=P\{X_1+X_2=1\mid H_1\}+P\{X_1+X_2=0\mid H_1\}=\frac{\lambda^k}{k!}e^{-\lambda}\Big|_{\lambda=2,k=1}+\frac{\lambda^k}{k!}e^{-\lambda}\Big|_{\lambda=2,k=0}=\frac{2}{e^2}+\frac{1}{e^2}=\frac{3}{e^3}.$

【例2】对正太总体的数学期望μ进行假设检验，如果在显著性水平$\alpha=0.05$下接受H_0：$\mu=\mu_0$，H_1：$\mu>\mu_0$，那么在$\alpha=0.01$下( )

(A)必然接受H_0 (B)必然拒绝H_0，接受H_1 (C)可能接受也可能拒绝H_0 (D)拒绝H_0，可能接受也可能拒绝H_1

【分析】如果是H_0：$\mu=\mu_0$，H_1：$\mu\neq\mu_0$，则是双边假设，现在是H_0：$\mu=\mu_0$，H_1：$\mu>\mu_0$，因此你可以写成

H_0：$\mu\leqslant\mu_0$，H_1：$\mu>\mu_0$，所以我就把我接受的范围向右开拓一个$\triangle$，我的接收域就是$(-\infty,\triangle)$，

题目没说σ已知还是未知，因此$\triangle$可能是$\mu_0+z_\alpha\frac{\sigma}{\sqrt{n}}$，也可能是$\mu_0+t_\alpha(n-1)\frac{S}{\sqrt{n}}$，本来我在$\alpha=0.05$的条件下，是接受$H_0$的

现在把α减小到了0.01，那么z_α和$t_\alpha(n-1)$，肯定是变大的，因为α是上α分位数，是右侧的面积，现在右侧的面积从0.05到0.01变小了

因此z_α和$t_\alpha(n-1)$会增大，所以我的接收域的区间变大了，包含了我之前的接收域，因此α减小后，仍然接受H_0。选A.

【区间估计^2】：总体X的分布已知，但是含有参数θ，其中$X_1\cdots X_n$是从中抽取的样本，观测值为$x_1\cdots x_n$，α是小概率，叫显著性水平，$1-\alpha$叫置信度

则$P\{\theta\in(a,b)\}=1-\alpha$，这个区间$(a,b)$叫$\theta$的置信度为$1-\alpha$的置信区间。

【注】考试中的总体一定是正太总体$X\sim N(\mu,\sigma^2)$：

一、μ的区间估计

①σ已知：用$\dfrac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\sim N(0,1)$去估计 $\Rightarrow P\left\{-z_{\frac{\alpha}{2}}<\dfrac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}<z_{\frac{\alpha}{2}}\right\}=1-\alpha\Rightarrow \mu\in\left(\overline{X}-z_{\frac{\alpha}{2}}\dfrac{\sigma}{\sqrt{n}},\ \overline{X}+z_{\frac{\alpha}{2}}\dfrac{\sigma}{\sqrt{n}}\right)$即$\mu$的置信度为$1-\alpha$的置信区间

②σ未知：用$\dfrac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}}\sim t(n-1)$去估计 $\Rightarrow P\left\{t_{\frac{\alpha}{2}}(n-1)<\dfrac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}}<t_{\frac{\alpha}{2}}(n-1)\right\}=1-\alpha\Rightarrow \mu\in\left(\overline{X}-t_{\frac{\alpha}{2}}(n-1)\dfrac{S}{\sqrt{n}},\ \overline{X}+t_{\frac{\alpha}{2}}(n-1)\dfrac{S}{\sqrt{n}}\right)$

二、σ^2的区间估计

①μ已知：用$\sum\limits_{i=1}^{n}\left(\dfrac{X_i-\mu}{\sigma}\right)^2\sim\chi^2(n)$去估计 $\Rightarrow P\left\{\chi^2_{1-\frac{\alpha}{2}}(n)<\sum\limits_{i=1}^{n}\left(\dfrac{X_i-\mu}{\sigma}\right)^2<\chi^2_{\frac{\alpha}{2}}(n)\right\}=1-\alpha\Rightarrow\sigma^2\in\left(\dfrac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{\frac{\alpha}{2}}(n)},\ \dfrac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{1-\frac{\alpha}{2}}(n)}\right)$

②μ未知：

用$\sum\limits_{i=1}^{n}\left(\dfrac{X_i-\overline{X}}{\sigma}\right)^2\sim\chi^2(n-1)=\dfrac{(n-1)S^2}{\sigma^2}$去估计 $\Rightarrow P\left\{\chi^2_{1-\frac{\alpha}{2}}(n-1)<\dfrac{(n-1)S^2}{\sigma^2}<\chi^2_{\frac{\alpha}{2}}(n-1)\right\}=1-\alpha\Rightarrow\sigma^2\in\left(\dfrac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)},\dfrac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}\right)$

【假设检验】还是正太总体$X\sim N(\mu,\sigma^2)$，样本X_1，$X_2\cdots X_n$来自X，x_1，$x_2\cdots x_n$为样本观测值，α叫显著性水平

一、对μ进行假设检验（还有对σ^2进行假设检验，但是不做整理）

①σ已知：
1、提出假设：H_0：$\mu=\mu_0$，H_1：$\mu\neq\mu_0$，（H_0是原始假设，H_1是备选假设，$\mu=\mu_0$表示μ和μ_0之间的差距不具有显著性）
2、$\dfrac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\sim N(0,1)\Rightarrow\left(-z_{\frac{\alpha}{2}},\ z_{\frac{\alpha}{2}}\right)$是原始假设$H_0$的接收域，
3、如果$\dfrac{\overline{x}-\mu_0}{\frac{\sigma}{\sqrt{n}}}\in\left(-z_{\frac{\alpha}{2}},\ z_{\frac{\alpha}{2}}\right)$，就接受$H_0$

【例】某次考试成绩服从$N(\mu,5^2)$，取36名考生成绩，已知$\overline{x}=66.5$，在显著性水平$\alpha=0.05$下，可否认为该批考试平均成绩为70分？

【分析】对μ进行假设检验，且σ已知：

1、设H_0：$\mu=70$，H_1：$\mu\neq70$；

2、$\dfrac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}=\dfrac{\overline{X}-\mu}{\frac{5}{6}}\sim N(0,1)$，且$\left(-z_{\frac{\alpha}{2}},\ z_{\frac{\alpha}{2}}\right)$为$\begin{matrix}(-z_{0.025},z_{0.025})\\(-1.96,1.96)\end{matrix}$，

上α分位数是$0.025=$右侧面积，所以查表看看$\Phi(x_0)=0.975$
x_0左侧的面积是0.975，这个x_0就是$z_{0.025}$
怎么查$\Phi(x_0)=0.975$：首先在正太分布表上找到0.975，
按行看表头值是1.9，看列表头值是0.06，所以$x_0=1.96$

3、现在看$\dfrac{\overline{x}-\mu_0}{\frac{\sigma}{\sqrt{n}}}=\dfrac{66.5-70}{\frac{5}{6}}=-4.2\notin(-1.96,1.96)$，所以拒绝$H_0$，即不认为平均成绩在70分.

②σ未知：
1、提出假设：$H_0=\mu_0$，H_1：$\mu\neq\mu_0$
2、$\dfrac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}}\sim t(n-1)\Rightarrow\left(-t_{\frac{\alpha}{2}}(n-1),\ t_{\frac{\alpha}{2}}(n-1)\right)$为$H_0$的接收域
3、看看$\dfrac{\overline{x}-\mu_0}{\frac{S}{\sqrt{n}}}$是否$\in\left(-t_{\frac{\alpha}{2}}(n-1),\ t_{\frac{\alpha}{2}}(n-1)\right)$，属于就接受$H_0$