概率论与数理统计B往年考题选解

任课教师: 王六权(wanglq@whu.edu.cn)

以下给出部分往年考题的解答, 多数是同学们问的比较多的题目, 也有部分是历年考试常考的题型. 如果解答中有错误, 请大家不吝指正.

1. (2016) 若随机变量(X, Y)的联合概率密度为

$$f(x,y) = \begin{cases} ae^{-\frac{1}{2}(x^2+y^2)} & x>0, y>0 \\ 0 & \text{其它} \end{cases}$$

(1) 求随机变量X和Y的边缘概率密度$f_X(x), f_Y(y)$. (2) X和Y是否独立? (3) 求$Z = X^2 + Y^2$的概率密度.

解: (1) 容易看出当$x < 0$时, $f_X(x) = 0$. 当$x > 0$时，我们有

$$f_X(x) = \int_0^\infty ae^{-\frac{1}{2}(x^2+y^2)}dy = ae^{-\frac{1}{2}x^2}\int_0^\infty e^{-\frac{1}{2}y^2}dy = \frac{\sqrt{2\pi}}{2}ae^{-\frac{1}{2}x^2}.$$

注意到

$$1 = \int_0^\infty f_X(x)dx = \frac{\sqrt{2\pi}}{2}a\int_0^\infty e^{-\frac{1}{2}x^2}dx = \frac{\pi}{2}a$$

由此解出$a = \frac{2}{\pi}$. 因此$f_X(x) = \sqrt{\frac{2}{\pi}}e^{-\frac{1}{2}x^2}I_{\{x>0\}}(x)$, 其中$I_A(x)$为示性函数。

由x, y的对称性可知$f_Y(y) = \sqrt{\frac{2}{\pi}}e^{-\frac{1}{2}y^2}I_{\{y>0\}}$.

(2) 因为$f(x,y) = f_X(x)f_Y(y)$, 故X, Y独立.

(3) 用积分转化法(参考教材119页例3.6.8)。对任何有界连续函数$h(z)$, 此时$g(x,y) = \sqrt{x^2+y^2}$. 在下面作积分变量代换: $x = \rho\cos\theta$, $y = \rho\sin\theta$, $\rho \geq 0$,

$0 \le \theta < 2\pi$, 使得$\sqrt{x^2+y^2}=\rho$,

$$\begin{aligned}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h[g(x,y)]f(x,y)dxdy &= \int_0^{\infty}\int_0^{\infty} h(x^2+y^2)ae^{-\frac{1}{2}(x^2+y^2)}dxdy \\ &= a\int_0^{\infty}\left(\int_0^{\frac{\pi}{2}} h(\rho^2)\rho e^{-\frac{1}{2}\rho^2}d\theta\right)d\rho \quad (\text{令}u=\rho^2) \\ &= \frac{a}{2}\int_0^{\infty} h(u)e^{-\frac{1}{2}u}\left(\int_0^{\frac{\pi}{2}}1d\theta\right)du \\ &= \frac{1}{2}\int_0^{\infty} h(u)e^{-\frac{1}{2}u}du.\end{aligned}$$

这里在最后一步中我们用到了第(1)问求出的结果:$a=\frac{2}{\pi}$. 由上式立即可知Z的概率密度为

$$f_Z(z)=\begin{cases}\frac{1}{2}e^{-\frac{1}{2}z} & z>0\\ 0 & z\le 0\end{cases}$$

2. (2007) 据报道：某地发现一个铁矿石，取25个样本测试，发现品位的平均值为68.5，样本方差为6.25. 如果说品位大于65即为高品位矿石。问: 此矿石是不是高品位的? ($\alpha=0.05$) （假设铁矿石品位近似服从正态分布）已知: $t_{0.05}(25)=1.708$, $t_{0.05}(24)=1.712$, $t_{0.025}(25)=2.060$, $t_{0.025}(24)=2.064$.

解 记$\mu_0=65$, 此题是在总体方差未知时检验均值μ, 且为如下右侧检验:

$$H_0:\mu\le\mu_0 \leftrightarrow H_1:\mu>\mu_0.$$ □

取检验统计量$u=\frac{\overline{X}-\mu_0}{S/\sqrt{n}}$, 拒绝域为$W=\{u>t_\alpha(n-1)\}$, 其中$n=25$. 计算可知$u=\frac{68.5-65}{\sqrt{6.25}/\sqrt{25}}=7$, 又已知$t_{0.05}(24)=1.712$, 因为$u>t_{0.05}(24)$, 故拒绝原假设$H_0$, 即认为该矿石是高品位。

说明 此题所问中没有“显著”二字，大家可能会疑惑到底应该怎么取备择假设。有的同学可能会问：为什么H_0不取为$\mu>\mu_0$，H_1取为$\mu\le\mu_0$? 这里提供以下两点理由，供大家参考:

(1) 虽然矿石抽样的品位均值68.5大于65，但是我们不确定这是抽样的随机性导致的还是总体显著大于65. 一般来说，只有矿石品位显著高才值得去开采，所

以我们实际上要看的就是矿石品位是否是显著高的。这样就应该把显著高作为备择假设H_1.

(2) 反过来想，如果真的取成相反的假设: $H_0 : \mu > \mu_0 \leftrightarrow H_1 : \mu \le \mu_0$. 此时的拒绝域会是$W = \{u < -t_\alpha(n-1)\}$, 因为样本均值$\overline{x} = 68.5 > \mu_0$, 显然$u > 0$, 即不在拒绝域里，这样我们连$t_\alpha(n)$分位数的值都用不上了。当然这点理由比较牵强。

这里再次指出：从不同的出发点考虑，假设可能会有不一样的提法。如果数据就是在假设值附近，买矿的希望尽量不要把低品位判断成高品味，卖矿的希望尽量不要把高品味判断成低品位，假设检验的做法会保护原假设，犯第一类错误的可能性小，从买矿的角度出发把低品位当原假设，从卖矿的角度出发把高品味当原假设。在本题中，我们是站在买（采）矿者的角度考虑的。

3. (2007) 若$X_1, X_2, \cdots, X_{16}$ 是总体$N(0,4)$的样本。
(1) 求$X = X_1^2 + X_2^2 + \cdots + X_{16}^2$的数学期望和方差.
(2) 确定a, 使得$t = a\frac{X_1+X_2+X_3+X_4}{\sqrt{X_5^2+X_6^2+\cdots+X_{16}^2}}$服从$t$ 分布，并求k.

解 (1) $\frac{1}{4}X = \left(\frac{X_1}{2}\right)^2 + \cdots + \left(\frac{X_{16}}{2}\right)^2$, 因为$\frac{X_i}{2} \sim N(0,1)$, 令$Y = \frac{1}{4}X$，则$Y \sim \chi^2(16)$. 我们有$E(Y) = 16$, $D(Y) = 2 * 16 = 32$ (教材209页例6.2.2结论). 因此$E(X) = 4E(Y) = 64$, $D(X) = 16D(Y) = 512$.

(2) 对分子和分母变形，使它们服从已知的分布。首先有$X_1 + X_2 + X_3 + X_4 \sim N(0,16)$, 从而$\frac{1}{4}(X_1 + X_2 + X_3 + X_4) \sim N(0,1)$. 又$\frac{1}{4}(X_5^2 + \cdots + X_{16}^2) \sim N(0,12)$, 因此我们有

$$\frac{\frac{1}{4}(X_1 + X_2 + X_3 + X_4)}{\sqrt{\frac{1}{4}(X_5^2 + \cdots + X_{16}^2)/12}} \sim t(12).$$

因此$a = \sqrt{3}$, $k = 12$. □

4. (2014, 2017) 若随机变量(X,Y)的联合概率密度为

$$f(x,y) = \begin{cases} e^{-x-y} & x, y > 0 \\ 0 & \text{其他} \end{cases}$$

(1) 求随机变量X和Y的边缘概率密度$f_X(x)$和$f_Y(y)$, 并判断它们是否独立?
(2) 求$Z = X - Y$的概率密度.

解 (1) 易知当$x \le 0$时$f_X(x) = 0$. 当$x > 0$时，我们有

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy = \int_0^{\infty} e^{-x-y}dy = e^{-x}.$$

因此$f_X(x) = e^{-x}I_{\{x>0\}}(x)$, 其中$I_A(x)$为集合$A$上的示性函数。同理有$f_Y(y) = e^{-y}I_{\{y>0\}}(y)$。

因为$f(x,y) = f_X(x)f_Y(y)$, 所以X, Y独立.

(2) 由卷积分式，我们有

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_X(x)f_Y(x-z)dx \\ &= \int_{x>0\text{且}x>z} e^{-x}e^{z-x}dx = \int_{x>\max\{0,z\}} e^{z-2x}dx. \end{aligned}$$

当$z \le 0$时，

$$f_Z(z) = \int_0^{\infty} e^{z-2x}dx = \frac{1}{2}e^z.$$

当$z > 0$时，

$$f_Z(z) = \int_z^{\infty} e^{z-2x}dx = \frac{1}{2}e^{-z}.$$

所以$Z = X - Y$的概率密度是

$$f_Z(z) = \begin{cases} \frac{1}{2}e^z & z \le 0 \\ \frac{1}{2}e^{-z} & z > 0 \end{cases}$$ □

说明 2017年仅将第二问改为$Z = X + Y$, 即使应用如下卷积分式求Z的概率密度:

$$f_Z(z) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_X(x)f_Y(z-x)dx.$$

5. (2017) 若总体X的概率密度为$f(x) = \begin{cases} \frac{1}{2\theta} & -\theta \le x \le \theta \\ 0 & \text{其它} \end{cases}$ $X_1, X_2, \cdots, X_n$ 为样本。(1) 求θ^2 的矩估计，并判别是否无偏。
(2) 求θ 的极大似然估计，并判别是否无偏。
(3) 可否求θ 的一个无偏的矩估计。

解 (1) $E(X^2) = \int_{-\theta}^{\theta} \frac{1}{2\theta}x^2dx = \frac{\theta^2}{3}$. 令$E(X^2) = A_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$, 解得$\theta^2$ 之矩估计

为$\hat{\theta^2}=\frac{3}{n}\sum_{i=1}^n X_i^2$. 因为$E(\hat{\theta^2})=\frac{3}{n}\sum_{i=1}^n E(X_i^2)=3E(X^2)=\theta^2$, 故$\hat{\theta^2}$ 是θ^2的无偏矩估计。

(2) 似然函数$L(\theta)=\left(\frac{1}{2\theta}\right)^n$是$\theta>0$的单调递减函数。注意到$|X_i|\le\theta$，我们有$\theta\ge\max\{|X_1|,\cdots,|X_n|\}$. 因此$L(\theta)$ 在$\theta=\max\{|X_1|,\cdots,|X_n|\}$ 时取得最大值，即θ 的极大似然估计为$\hat{\theta}=\max\{|X_1|,\cdots,|X_n|\}$.

接下来先求$|X|$ 的分布。当$0\le t\le\theta$时我们有

$$P(|X|\le t)=P(-t\le X\le t)=\frac{t}{\theta}.$$

因此

$$\begin{aligned}P(\hat{\theta}\le t)=&P(\max\{|X_1|,\cdots,|X_n|\}\le t)=P(|X_1|\le t)\cdots P(|X_n|\le t)\\=&P(|X|\le t)^n\left(\frac{t}{\theta}\right)^n.\end{aligned}$$

于是$\hat{\theta}$ 的分布函数为

$$F(t)=\begin{cases}0 & t<0\\ \left(\frac{t}{\theta}\right)^n & 0\le t\le\theta\\ 1 & t\ge\theta\end{cases}$$

其概率密度为

$$f(t)=\begin{cases}0 & t<0\text{或}t>\theta\\ n\frac{t^{n-1}}{\theta^n} & 0\le t\le\theta\end{cases}$$ □

因此$E(\hat{\theta})=\int_0^\theta t\cdot n\frac{t^{n-1}}{\theta^n}dt=\frac{n\theta}{n+1}$. 因为$E(\hat{\theta})\ne\theta$, 故$\hat{\theta}$不是无偏的。

(3) 注意到$E(|X|)=2\int_0^\theta x\cdot\frac{1}{2\theta}dx=\frac{\theta}{2}$, 令$E(|X|)=\frac{1}{n}\sum_{i=1}^n|X_i|$, 解得$\theta$ 的一个矩估计为$\hat{\theta}_2=\frac{2}{n}\sum_{i=1}^n|X_i|$. 易知$E(\hat{\theta}_2)=\frac{2}{n}\sum_{i=1}^n E(|X_i|)=\theta$, 即$\hat{\theta}_2$ 为θ的一个无偏矩估计。

6. (2014) 甲、乙两个棋迷意外得到900元, 他们以下棋来决定这笔钱的归属：先赢三盘的人拿走全部的钱. 下完三盘后意外中止, 此时甲二胜一负, 乙说：你拿600, 我拿300.如果这是他们两个人的真实水平, 问：这个分法合理吗? 说明理由, 你可不可以给出一个更合理的分配方案?

解 每盘棋甲胜概率为$\frac{2}{3}$, 乙胜概率为$\frac{1}{3}$. 如果比赛继续进行, 那么只可能以三种方式结束, 即接下来胜的人和次序是甲, 乙甲或乙乙, 前两者都是甲赢（拿走全部钱）.

$$P(\text{甲赢}) = P(\text{甲}) + P(\text{乙甲}) = \frac{2}{3} + \frac{1}{3}\cdot\frac{2}{3} = \frac{8}{9}.$$

类似地, $P(\text{乙赢}) = \frac{1}{9}$. 故分配不合理. 应给甲800元, 乙100元. □

7. (2014) 若随机变量X在区间$(0,8)$服从均匀分布. (1) 求方程$y^2+2y+X=0$ 有实根的概率. (2) 若对随机变量X 进行4 次独立观察, 记Y 为上述方程有解的次数, 求Y 的数学期望和方差.

   解 (1) 方程有实根等价于$\Delta = 4-4X \geq 0$, 即$X \leq 1$, 概率$P(X \leq 1) = \frac{1}{8}$.

   (2) $Y \sim B(4, \frac{1}{8})$, $n=4$, $p=\frac{1}{8}$, $E(Y) = np = \frac{1}{2}$, $D(Y) = np(1-p) = \frac{7}{16}$. □

8. (2016) 某生产线加工产品的合格率为0.8, 已知: 合格每件可获利80 元, 不合格每件亏损20元.

   (1) 为保证每天的平均利润不低于6000, 问他们至少要加工多少件产品?

   (2) 为保证每天的利润不低于6000的概率大于0.977, 问他们至少要加工多少件产品? （已知$\Phi(2.0) = 0.977$).

   解 用X表示每件产品所获利润, 则$E(X) = 80\times 0.8 - 20\times 0.2 = 60$, $E(X^2) = 5200$, $D(X) = E(X^2) - (E(X))^2 = 1600$.

   (1) 设每天加工n件产品, 设第i 件产品所获利润为X_i. 则要求$E(\sum_{i=1}^n X_i) \geq 6000$, 即$nE(X) \geq 6000$, 故需要$n \geq 100$.

   (2) 要保证$P(\sum_{i=1}^n X_i \geq 6000) > 0.977$, 由中心极限定理, 需要

$$\frac{6000-60n}{40\sqrt{n}} < -2,$$

   解之得$\sqrt{n} > \frac{4+\sqrt{3616}}{6} = \frac{2+2\sqrt{226}}{3}$. 计算知$\sqrt{n} > 10.7$, 故$n > 114.5$, 即$n > 115$. □

9. (2018, 题七) 若$X_1, X_2, \ldots, X_n$是相互独立的随机变量, 而且都在区间$(-\theta,\theta)$服从均匀分布. 令

$$M = Max\{|X_1|, |X_2|, \ldots, |X_n|\}, \quad N = Min\{|X_1|, |X_2|, \ldots, |X_n|\}.$$

试求M,N的期望与方差.

解 设$X\sim U(-\theta,\theta)$, 下面求$|X|$的分布函数$F(x)=P(|X|\le x)$.

当$x<0$时, 显然$F(x)=0$. 当$x\ge\theta$时, $F(x)=1$. 当$0\le x<\theta$时, $F(x)=P(-x\le X\le x)=\frac{2x}{2\theta}=\frac{x}{\theta}$.

因此$|X|$的分布函数为

$$F(x)=\begin{cases}0 & x<0,\\ \frac{x}{\theta} & 0\le x<\theta,\\ 1 & x\ge\theta.\end{cases}$$

接下来求M的分布函数. 我们有

$$\begin{aligned}F_M(x)=P(M\le x)=P(|X_1|\le x)\cdots P(|X_n|\le x)=F(x)^n\\ =\begin{cases}0 & x<0,\\ \frac{x^n}{\theta^n} & 0\le x<\theta,\\ 1 & x\ge\theta.\end{cases}\end{aligned}$$

于是M的概率密度函数为

$$f_M(x)=\begin{cases}0 & x<0,\\ n\frac{x^{n-1}}{\theta^n} & 0\le x<\theta,\\ 0 & x\ge\theta.\end{cases}$$

于是

$$E(M)=\int_{-\infty}^{\infty}xf_M(x)dx=\int_0^{\theta}n\frac{x^n}{\theta^n}dx=\frac{n}{n+1}\theta.$$

$$E(M^2)=\int_{-\infty}^{\infty}x^2f_M(x)dx=\int_0^{\theta}n\frac{x^{n+1}}{\theta^n}dx=\frac{n}{n+2}\theta^2.$$

于是

$$D(M)=E(M^2)-(E(M))^2=\frac{n}{(n+1)^2(n+2)}\theta^2.$$

(2) 先求N的分布函数,

$$\begin{aligned}F_N(x) &= P(N \le x) = 1 - P(N > x) \\ &= 1 - P(X_1 > x)\cdots P(X_n > x) = 1 - (1 - F(x))^n \\ &= \begin{cases} 0 & x < 0, \\ 1 - (1 - \frac{x}{\theta})^n & 0 \le x < \theta, \\ 1 & x \ge \theta. \end{cases}\end{aligned}$$

□

于是N的概率密度函数为

$$f_N(x) = \begin{cases} 0 & x < 0, \\ \frac{n}{\theta}(1 - \frac{x}{\theta})^{n-1} & 0 \le x < \theta, \\ 0 & x \ge \theta. \end{cases}$$

同上可求得

$$E(N) = \int_0^\theta n\frac{x}{\theta}(1 - \frac{x}{\theta})^{n-1}dx = \frac{1}{n+1}\theta,$$

$$E(N^2) = \int_0^\theta n\frac{x^2}{\theta}(1 - \frac{x}{\theta})^{n-1}dx = \frac{2}{(n+1)(n+2)}\theta^2.$$

$$D(N) = E(N^2) - (E(N))^2 = \frac{n}{(n+1)^2(n+2)}\theta^2.$$

10. (2018, 题八) 若随机变量(X, Y)服从二维正态分布, 它们的联合概率密度为

$$f(x, y) = \frac{1}{2\pi}\exp\{-\frac{1}{2}(x^2 + y^2)\}.$$

(a) 求随机变量$Z = 2X - Y$的概率密度;

(b) 求Z与$W = 2X + Y$的相关系数.

解 (1) 设$(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, 由其联合分布函数可见$\mu_1 = \mu_2 = 0$, $\sigma_1 = \sigma = 1$, $\rho = 0$. 因$\rho = 0$, 故X, Y 独立, 并且$X \sim N(0, 1), Y \sim N(0, 1)$. 由正态分布的再生性, 可知$Z = 2X - Y$仍服从正态分布. 又$E(Z) = 2E(X) - E(Y) = 0$, $D(Z) = D(2X - Y) = D(2X) + D(-Y) = 4D(X) + D(Y) = 5$, 故$Z \sim N(0, 5)$. 因此$Z \sim N(0, 5)$, 从而$f_Z(z) = \frac{1}{\sqrt{10\pi}}\exp\{-\frac{1}{10}z^2\}$.

(2) 同上可知$W \sim N(0,5)$. 我们有

$$\begin{aligned}\mathrm{Cov}(Z,W) &= \mathrm{Cov}(2X-Y,2X+Y)\\ &= \mathrm{Cov}(2X,2X)+\mathrm{Cov}(2X,Y)+\mathrm{Cov}(-Y,2X)+\mathrm{Cov}(-Y,Y)\\ &= 4\mathrm{Cov}(X,X)+2\mathrm{Cov}(X,Y)-2\mathrm{Cov}(Y,X)-\mathrm{Cov}(Y,Y)\\ &= 4D(X)-D(Y)=3. \qquad \square\end{aligned}$$

因此

$$\rho_{ZW}=\frac{Cov(Z,W)}{\sqrt{D(Z)}\sqrt{D(W)}}=\frac{3}{5}.$$

11. (2016) 若总体X在$(\theta,1)$ 上服从均匀分布, θ 未知, $X_1,X_2,\dots,X_n$ 为样本.
(1) 求θ 的矩估计; (2) 求θ 的极大似然估计；(3) 它们是否为无偏估计, 如果不是, 将它们化为无偏估计.

解 (1) 首先有$E(X)=\frac{\theta+1}{2}$, 令$E(X)=\overline{X}$, 得θ 之矩估计为$\hat{\theta}_1=2\overline{X}-1$.

(2) 似然函数

$$L(\theta)=\left(\frac{1}{1-\theta}\right)^n.$$

注意到$L(\theta)$是θ 的单调增函数. 因为$\theta \le X_i \le 1$ $(i=1,2,\dots,n)$, 我们有$\theta \le \min\{X_1,\dots,X_n\}$. 故$L(\theta)$ 在$\theta=\min\{X_1,X_2,\dots,X_n\}$ 时取得最大值. 从而θ 的极大似然估计量是$\hat{\theta}_2=\min\{X_1,X_2,\dots,X_n\}$.

(3) 因为$E(\hat{\theta}_1)=2E(\overline{X})-1=2E(X)-1=\theta$, 故$\hat{\theta}_1$ 是θ 的无偏估计.

注意到

$$P(\hat{\theta}_2>t)=P(X_1>t)\cdots P(X_n>t)=\left(\frac{1-t}{1-\theta}\right)^n.$$

因此$\hat{\theta}_2$的分布函数为

$$F(t)=P(\hat{\theta}_2\le t)=1-\left(\frac{1-t}{1-\theta}\right)^n.$$

其概率密度函数为

$$f(t)=\begin{cases} n\frac{(1-t)^{n-1}}{(1-\theta)^n}, & \theta\le t\le 1. \\ 0 & \text{其他}\end{cases}$$

因此

$$E(\hat{\theta}_2)=\int_\theta^1 nt\frac{(1-t)^{n-1}}{(1-\theta)^n}dt=\frac{1+n\theta}{n+1}.$$

因为$E(\hat{\theta}_2)\neq\theta$, 故$\hat{\theta}_2$不是无偏估计.

令$\hat{\theta}_3=\frac{n+1}{n}(\hat{\theta}_2-\frac{1}{n+1})$, 则$\hat{\theta}_3$是$\theta$的无偏估计. □

12. (2017) 若$X_1,X_2,\cdots,X_n$是正态总体$N(0,4)$的样本, $\overline{X}=\frac{1}{n}\sum_{i=1}^n X_i$, $S^2=\frac{1}{n}\sum_{i=1}^n X_i^2$. (1) 求$\overline{X},S^2$ 的数学期望和方差. (2) 确定k, 使得$t=k\frac{\overline{X}}{S}$ 服从t分布.

解 (1) $E(\overline{X})=E(X)=0$, $D(\overline{X})=\frac{4}{n}$. 注意到$E(X^2)=D(X)+(E(X))^2=4+0=4$, 因此

$$E(S^2)=\frac{1}{n}\sum_{i=1}^n E(X_i^2)=E(X^2)=4.$$

因为$E(X^4)=2^4(4-1)!!=48$ (教材170页例4.4.1), 故$D(X^2)=E(X^4)-(E(X^2))^2=32$. 故$D(S^2)=\frac{1}{n^2}nD(X^2)=\frac{32}{n}$.

(2) $\overline{X}\sim N(0,\frac{4}{n})$, 从而$\frac{\sqrt{n}}{2}\overline{X}\sim N(0,1)$. 注意到$S^2=\frac{4}{n}\sum_{i=1}^n\left(\frac{X_i}{2}\right)^2$, 因为$\frac{X_i}{2}\sim N(0,1)$, 故$\frac{n}{4}S^2\sim\chi^2(n)$. 于是有

$$\frac{\frac{\sqrt{n}}{2}\overline{X}}{\sqrt{\frac{n}{4}S^2/n}}\sim t(n).$$

故应取$k=\sqrt{n}$. □

注意 上题中的S^2**并非**课本里学的样本方差, 这是一个陷阱.

13. (2016) 某地发现一个金属矿, 取25个样本测试, 发现品位的平均值为32.1, 样本方差为6.25. 问：此矿的品位是不是显著大于30? ($\alpha=0.05$) （假设矿石

品位近似服从正态分布）已知: $u_{0.05}=1.65, u_{0.025}=1.96$, $t_{0.05}(25)=1.708$, $t_{0.05}(24)=1.712$, $t_{0.025}(25)=2.060$, $t_{0.025}(24)=2.064$.

说明: 这里的u_α即为z_α, 即正态分布的上α分位数.

解 此为σ^2未知关于μ的假设检验, 且为如下右侧检验:

$$H_0:\mu=\mu_0 \leftrightarrow H_1:\mu>\mu_0,$$

其中$\mu_0=30$. 取检验统计量$T=\frac{\overline{X}-\mu_0}{S/\sqrt{n}}$, 拒绝域$W=\{T>t_\alpha(n-1)\}$. 注意到$n=25$, $t_{0.05}(24)=1.712$, $\overline{x}=32.1$, $S=\sqrt{6.25}=2.5$, 代入得$T=4.2$, 因为$T>1.712$, 故拒绝H_0. □

注意 (1) 上面的假设不能取错！因为所问有显著二字, 故将大于30取为备择假设H_1. (2) 对于H_0的选取则自由一些，上面是根据经验，μ 应当和μ_0很接近，所以将H_0取为$\mu=\mu_0$. 但也可以选为$H_0:\mu\le\mu_0$。

14. (2014) 若总体X 服从正态分布$N(\mu,1)$, $X_1,X_2,\cdots,X_n$ 为样本.
(1) 若想要μ 的0.95的置信区间长度小于0.5, 样本容量n至少要多大?
(2) 若某次取样, $n=25$, $\overline{X}=76.5$, 可否认为μ显著大于76? ($\alpha=0.05$) ($z_{0.05}=1.65$, $z_{0.025}=1.96$).

解 (1) 这是$\sigma^2=1$ 已知时对μ 的区间估计, 回忆μ 的置信度为$1-\alpha$ 的置信区间为$\left(\overline{X}-\frac{\sigma}{\sqrt{n}}z_{\frac{\alpha}{2}},\overline{X}+\frac{\sigma}{\sqrt{n}}z_{\frac{\alpha}{2}}\right)$, 其长度为$2\frac{\sigma}{\sqrt{n}}z_{\frac{\alpha}{2}}$. 取$\alpha=0.05$, 要$2\frac{\sigma}{\sqrt{n}}z_{\frac{\alpha}{2}}<0.5$, 需有$n\ge 61.4$, 故$n\ge 62$.

(2) 令$\mu_0=76$. 检验假设$H_0:\mu\le\mu_0 \leftrightarrow H_1:\mu>\mu_0$. 统计量为$u=\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$, 拒绝域为$W=\{u>z_\alpha\}$. 计算得$u=2.5>1.65$. 故拒绝$H_0$, 即可以认为$\mu$显著大于76. □