第六章习题解析

1. 利用单样本均值估算总体均值置信区间时，由于$\frac{\bar{x}-\mu}{s/\sqrt{n}}$符合自由度为n-1的t分布。
   于是对于μ的$1-\alpha$的置信区间有$-t_{\alpha/2,n-1}<\frac{\bar{x}-\mu}{s/\sqrt{n}}<t_{\alpha/2,n-1}$，
   于是有$\bar{x}-t_{\alpha/2,n-1}\times s/\sqrt{n}<\mu<\bar{x}+t_{\alpha/2,n-1}\times s/\sqrt{n}$，将数据代入可得：
   解：
   下限：$\bar{x}-t_{\alpha/2,n-1}\times s/\sqrt{n}=103-2.06\times 2/5=102.31$
   上限：$\bar{x}+t_{\alpha/2,n-1}\times s/\sqrt{n}=103+2.06\times 2/5=103.69$
2. 同前，$\frac{\bar{x}-\mu}{s/\sqrt{n}}$符合自由度为n-1的t分布。
   于是对于μ的$1-\alpha$的置信区间有$-t_{\alpha/2,n-1}<\frac{\bar{x}-\mu}{s/\sqrt{n}}<t_{\alpha/2,n-1}$
   题中没有给出s，给出的是$\sum x^2$，可以借助$s^2=\frac{\sum x^2-n\bar{x}^2}{n-1}$计算s。
   解：
   $s=\sqrt{\frac{\sum x^2-n\bar{x}^2}{n-1}}=\sqrt{\frac{265317-25\times 103^2}{24}}=1.96$
   下限：$\bar{x}-t_{\alpha/2,n-1}\times s/\sqrt{n}=103-2.06\times 1.96/5=102.19$
   上限：$\bar{x}+t_{\alpha/2,n-1}\times s/\sqrt{n}=103+2.06\times 1.96/5=103.81$
3. 当2样本率数据满足：$n_1p_1>5,n_1(1-p_1)>5,n_2p_2>5,n_2(1-p_2)>5$的时候，2样本有效率差值可以近似符合正态分布，作标准正态分布转化有：
   $-z_{\alpha/2}<\frac{(p_1-p_2)-(\pi_1-\pi_2)}{s_{p_1-p_2}}<z_{\alpha/2}$
   其中π_1,π_2分别表示样本1来源总体和样本2来源总体的有效率
   而$s_{p_1-p_2}=\sqrt{p_c(1-p_c)(\frac{1}{n_1}+\frac{1}{n_2})},p_c=\frac{x_1+x_2}{n_1+n_2}$
   解：
   由于$n_1p_1=68,n_1(1-p_1)=12,n_2p_2=64,n_2(1-p_2)=19$均大于5，可以认为2样本率的差近似服从正态分布。
   $p_c=\frac{68+64}{80+83}=0.79$
   $s_{p_1-p_2}=\sqrt{p_c(1-p_c)(\frac{1}{n_1}+\frac{1}{n_2})}=\sqrt{0.79\times(1-0.79)(\frac{1}{80}+\frac{1}{84})}=0.06$
   下限：$p_1-p_2-z_{\alpha/2}\times s_{p_1-p_2}=0.08-1.96*0.06=-0.04$
   上限：$p_1-p_2+z_{\alpha/2}\times s_{p_1-p_2}=0.08+1.96*0.06=0.20$
4.

问题1.
(1) 建立检验假设，确定检验水准
$H_0:\mu=\mu_0$，树苗成长5年后平均高度为1.8米
$H_1:\mu\neq\mu_0$，树苗成长5年后平均高度不是1.8米
$\alpha=0.01$
(2) 计算检验统计量
已知总体$\mu_0=1.8$，标准差$\sigma=0.3$，在H_0成立的假设下，样本均值$\bar{x}$~$N(\mu_0,\frac{\sigma^2}{n})$
可得$z=\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}=\frac{1.6-1.8}{0.3/9}=-6$

(3) 根据p值做出推断结论
α=0.01, 可得$z_{\alpha/2} = 2.58$, 由于z=-6<-2.58，可得P<0.05, 差别具有统计学意义，拒绝H_0，接受H_1，认为树苗成长5年后平均高度不是1.8米

问题2.
在前面已经说明，$z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$符合$N(0,1)$的标准正态分布。于是合格标准的下限z及上限z分别是
下限：$z_1 = \frac{1.73-1.8}{0.3/9} = -2.1$
上限：$z_2 = \frac{1.87-1.8}{0.3/9} = 2.1$
通过查表可得标准正态分布曲线z1左侧及z2右侧的面积均为0.02, 这两个面积合起来就是在这一评判标准下认定不合格（假设检验阳性）但实际上树苗合格（样本实际来自合格的总体，实际为阴性）的概率，即I类错误概率：0.04

问题3.
根据问题3，树苗实际来源于成长5年后，树苗高度均值为1.5米，标准差0.4米的总体，于是有$z = \frac{\bar{x}-\mu 1}{\sigma 1/\sqrt{n}} = \frac{\bar{x}-1.5}{0.4/\sqrt{81}}$符合N(0,1)的正态分布那么在$\bar{x}$=1.73米和$\bar{x}$=1.87米的两个临界值之间（假设检验阴性），由新的μ_1, σ_1得出的$z = \frac{\bar{x}-\mu 1}{\sigma 1/\sqrt{n}}$标准正态分布的曲线（样本实际来自不合格的总体，实际为阳性）下面积就是II类错误的概率。
由不合格总体参数，按照当前认定合格的标准可得在新的标准正态分布曲线下，下限与上限分别是：
下限：$z_1 = \frac{1.73-1.5}{0.4/9} = 5.18$
上限：$z_2 = \frac{1.87-1.5}{0.4/9} = 8.33$
两者正态曲线下面积均显著小于0.001，其差值也小于0.001