陈思杰
2017280625

西南交通大学在职研究生应用统计作业

一、选择题

1、某家具制造商购买大批木材，木材不干会影响家具的尺寸和形状。家具制造商从每批货中随机抽取 5 块木材检验湿度，如果其中任何一块木材的湿度超过标准，就把整批货退回。这个问题中（B D）。

A、样本是从所有木材批次中随机抽取的部分批次木材；

B、样本是从每批木材中随机抽取的 5 块木材；

C、总体单位是从所有木材批次中随机抽取的部分批次木材；

D、总体是购买的全部木材。

2、设有一个直线回归方程为 $y=2-2.5x$，则变量 x 增加一个单位时（C）

A、y 平均增加 2.5 个单位；　　B、y 平均增加 2 个单位；

C、y 平均减少 2.5 个单位；　　D、y 平均减少 2 个单位。

3、若两个变量的平均水平接近，标准差越大的变量，其（ B ）。

A、标准差代表性越大；　　B、离散程度越大；

C、稳定性越高；　　D、分布偏斜程度越严重。

4、某研究部门准备在全市 200 万个家庭中抽取 2000 个家庭，据此推断该城市所有职工家庭的年人均收入，这项研究的统计量是（C）。

A、2000 个家庭；　　B、200 万个家庭；

C、2000 个家庭的人均收入；　　D、200 万个家庭的人均收入。

5、P 值所反映的是（D）。

A、拒绝域的大小；

B、统计量的大小；

C、事先给定的显著性水平的大小；

D、若原假设为真，所得到的样本结果会像实际观测结果那么极端或更极端的概率。

6、对于右偏分布，均值、中位数和众数之间的关系是（A）。

A、均值>中位数>众数；　　B、. 中位数>均值>众数；

C、众数>中位数>均值；　　D、 众数>均值>中位数。

7、设总体 $X\sim N(\mu,\sigma^2)$，其中 σ^2 已知，若已知样本容量 n 和置信度 $1-\alpha$ 均不变，则对于不同的样本观察值，总体均值 μ 的置信区间的长度（C）。

A、变长　B、变短　C、不变　D、不能确定

8、在假设检验中，显著性水平β的意义是（A）。

A、H_1为真，但经检验拒绝H_1的概率；　　B、H_1为真，经检验接受H_1的概率；

C、H_1不成立，经检验拒绝H_1的概率；　　D、H_1不成立，但经检验接受H_1的概率。

9、在假设检验中，用α和β分别表示犯第一类错误和第二类错误的概率，则当样本容量一定时，下列说法正确的是（C）。

A、α减少，β也减少；　　B、α增大，β也增大

C、α与β不能同时减少，其中一个减少，另一个往往会增大；　D、A和B同时成立。

10、根据判定系数与F统计量的关系可知，当R^2=1时，有（D）。

A、F=1；　　B、F= —1；　　C、F=0；　　D、F=∞

二、计算题

1、一个求职者参加了两个公司的能力测试。在A公司的测试中，其平均分数是90分，标准差是10分；在B公司测试中，其平均分数是120分，标准差是25分。此人在A公司的测试中得了115分，在B公司的项测试中得了170分。与平均分数相比，此人在哪一个公司测试的成绩更为理想？（5分）

$$A=\frac{10}{90}=11.11\%$$

$$B=\frac{25}{110}=20.83\%.$$

2、某城市环保局在该城市随机选取10个测量点测量了24小时空气污染物质量(μg/m3)，样本均值为88μg/m3，样本标准差为7.5μg/m3。假设污染物质量服从正态分布，试推测该城市24小时平均污染物质量的置信区间（1-α=0.95）（10分）

$$X\sim N(88,7.5^2)$$

置信区间 $\left(\mu-z(\frac{\alpha}{2})\cdot\frac{\sigma}{\sqrt{n}},\ \mu+z(\frac{\alpha}{2})\cdot\frac{\sigma}{\sqrt{n}}\right)$

$$=\left(88-z(\tfrac{\alpha}{2})\cdot\frac{7.5}{\sqrt{16}},\ 88+z(\tfrac{\alpha}{2})\cdot\frac{7.5}{\sqrt{16}}\right)$$

$$=\left(88-1.96\cdot\frac{7.5}{4},\ 88+1.96\cdot\frac{7.5}{4}\right)=(84.325,\ 91.675)$$

3、调节一个装瓶机使其对每个瓶子的灌装量均值为μ盎司，通过观察这台装瓶机对每个瓶子的灌装量服从标准差$\sigma=1.0$盎司的正态分布。随机抽取由这台机器灌装的瓶子形成一个样本，并测定每个瓶子的灌装量。如果我们希望样本均值$\bar{X}$与μ的偏差在 0.3 盎司之内的概率达到 0.95，应当抽取多大的样本？（10 分）

$$P(|\bar{X}-\mu|\le 0.3)=P\left(\frac{|\bar{X}-\mu|}{\sigma/\sqrt{n}}\le\frac{0.3}{\sigma/\sqrt{n}}\right)$$

$$=P\left(\frac{-0.3}{1/\sqrt{n}}\le\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\le\frac{0.3}{1/\sqrt{n}}\right).$$

$$=2\phi(0.3\sqrt{n})-1\ge 0.95 \Rightarrow \phi(0.3\sqrt{n})\ge 0.975.$$

$$\Rightarrow 0.3\sqrt{n}\ge 1.96 \Rightarrow n\ge 42.6828 \quad \therefore n\ge 43.$$

4、一个农业试验站想要比较两个新的玉米品种的产量。研究人员认为不同农场之间的产量的差异可能会很大，故在 7 个农场的每一个农场都选择了两块一英亩的土地随机播种这两种玉米。成熟时收获玉米，试验结果（以蒲式耳为单位）如下表所示：

农场	1	2	3	4	5	6	7	合计	平均	标准差
品种 A	48.2	44.6	49.7	40.5	54.6	47.1	51.4	336.1	48.01429	4.594717
品种 B	41.5	40.1	44	41.2	49.8	41.7	46.8	305.1	43.58571	3.532435
差值	6.7	4.5	5.7	-0.7	4.8	5.4	4.6	31	4.428571	2.387268

问：根据这些数据检验两种玉米平均产量是否存在差异。（α=5%）（10 分）

$t_{\frac{\alpha}{2}}\,\sigma/\sqrt{n}$.

置信区间 A：$[48.01429-4.249405,\ 48.01429+4.249405]$.

B：$[43.58571-3.266957,\ 43.58571+3.266957]$.

$\therefore$ 存在差异.

5、某种新轮胎进行耐磨试验，资料如下。

试验小时数（小时）	13	25	27	46	18	31	46	57	75	87
磨损程度（系数）	0.1	0.2	0.2	0.3	0.1	0.2	0.3	0.4	0.5	0.6

采用 excel 计算得到如下结果：

	试验小时数（小时）	磨损程度（系数）
平均	42.5	0.29
标准差	24.54135	0.166333
方差	602.2778	0.027667
求和	425	2.9
观测数	10	10
$\sum(x-\bar{x})(y-\bar{y})$	36.55	

方差分析

	平方和
回归分析	0.246454
残差	0.002546
总计	0.249

	Coefficients	标准误差	t Stat	P-value
Intercept	0.003426	0.011743	0.291749	0.777901
试验小时数（小时）	0.006743	0.000242	27.82668	3E-09

要求：（1）试验小时数和磨损程度线性相关关系是否显著？（5 分）

（2）建立线性回归方程，用试验小时数估计磨损程度。（5 分）

（3）检验回归方程的线性关系（α=0.05）。（要求检验过程）（5 分）

（4）试验小时数为 50 小时，95%的置信水平下，轮胎磨损的预测区间是多少？（5 分）

6、管理人员对三位技术人员在四台不同机器上操作三天的生产率感兴趣，分别记录了这三天的日产量：

机器型号	技术人员								
	甲			乙			丙		
A	15	15	17	19	19	16	16	18	21
B	17	17	17	15	15	15	19	22	22
C	15	17	16	18	17	16	18	18	18
D	18	20	22	15	16	17	17	17	17

经分析得到如下方差分析表：

方差分析					
因变量: 生产率					
源	平方和	自由度	均方	F	P 值
技术人员	27.17	2	13.583	<1	0.002
机器	2.750	3	0.92	7.9	0.665
技术人员 * 机器	73.5	6	12.25	7.12	0.000
误差	41.333	24.	1.72	——	——
总计	144.750	35	——	——	——

要求：（1）将方差分析表中所缺的数值补齐；（6 分）

（2）分析技术人员、机器以及技术人员与机器的交互作用对生产率是否有显著影响（α=0.05）（要求检验过程）（9 分）

∵ $F_B = 7.9 > F_{0.95}(2,24) = 3.4$.

$F_{A\times B} = 7.12 > F_{0.95}(6.24) = 2.51$.

∴在α=0.05水平下.技术人员有显著差异.机器之间无显著差异.
交互作用有显著差异.