3.8 假设检验 参考答案

因为本次作业同学们犯的错误比较一致，所以先总结在前面。

总结：

1. 这两部分习题解题方法是很相似、有套路的，批改作业的时候发现一些同学掌握得还不是很好，要多加总结练习，不要被题目吓住，分析题目中给出的条件，代入对应的方法套路中。

2. 注意按照题目应该采取的是双侧检验还是单侧检验，是用哪种检验模型。

3. 注意计算问题，很多同学明明过程都是正确的，偏偏最后算出来的数是错的，这个一定要注意，不要白白丢分。

4. 注意解题步骤书写规范，注意适当穿插文字，不要整道题全是字母、数字，一个汉字都看不到，要适当加入说明性文字，把整道题像小作文一样书写下来。

5. 注意语言的规范，假设检验题目很多在最后给出结论时要写“有 xx%的把握认为怎样怎样”，不要直接写“不行”“不能”类似的答案。

3.8.1 已知某炼铁厂铁水含碳量服从正态分布$N(4.55, 0.108^2)$，现在测定了 9 炉铁水，其平均含碳量为 4.484，如果方差没有变化，可否认为现在生产之铁水平均含碳量仍为 $4.55(\alpha = 0.05)$？

解 设贴水含碳量 $X \sim N(\mu, 0.108^2)$，$\bar{x} = 4.484$，$n = 9$.

假设 $H_0: \mu = 4.55 \quad H_1: \mu \neq 4.55$

因为总体方差已知，选取统计量$U = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$

当H_0成立时，$U = \frac{\bar{X}-4.55}{\sigma/\sqrt{n}} \sim N(0,1)$

对于给定的显著性水平$\alpha = 0.05$，所作检验的模型为模型（Ⅰ），拒绝域为

$\{|u_0| \geq U_{\frac{\alpha}{2}}\}$，即$P(-U_{\frac{\alpha}{2}} < U < U_{\frac{\alpha}{2}}) = \Phi(U_{\frac{\alpha}{2}}) - \Phi(-U_{\frac{\alpha}{2}}) = 1 - \alpha = 0.95$

所以 $\Phi(U_{\frac{\alpha}{2}}) = 1 - \frac{\alpha}{2} = 0.975$

查表得临界值$U_{\frac{\alpha}{2}} = U_{0.025} = 1.96$

根据样本观察值，计算统计观测值

$$|u_0| = |\frac{\bar{x} - 4.55}{\sigma/\sqrt{n}}| = |\frac{4.484 - 4.55}{0.108/\sqrt{9}}| = 1.833 < U_{0.025} = 1.96$$

所以统计量没有落在拒绝域内，即有 95%的把握认为现在生产之铁水平均含碳量仍为 4.55。

注：本题正确率比较高，部分同学书写不够规范，也有计算错误。

3.8.3 一种原件，要求其使用寿命不得低于 1000(h)。现在从这一批这种原件中随机抽取 25 件，测得其寿命平均值 950 (h)。已知这种元件寿命服从标准差 $\sigma = 100$(h)的正态分布，试在显著性水平 $\alpha = 0.05$ 下确定这批原件是否合格。

解　设元件寿命$X \sim N(\mu, 0.108^2)$，$\overline{x} = 950$，$n = 25$.

假设 $H_0: \mu \geq 1000$　$H_1: \mu \neq 1000$

与模型 $H_0: \mu = 1000$　$H_1: \mu \neq 1000$ 有相同拒绝域

因为总体方差已知，选取统计量$U = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$

当H_0成立时，$U = \frac{\overline{X}-1000}{\sigma/\sqrt{n}} \sim N(0,1)$

对于给定的显著性水平$\alpha = 0.05$，所作检验的模型为模型（III），拒绝域为$\{u_0 \leq -U_\alpha\}$，即

$$P(U < -U_\alpha) = \Phi(-U_\alpha) = 1 - \Phi(U_\alpha) = \alpha = 0.05$$

所以$\Phi(U_\alpha) = 1 - \alpha = 0.95$，

查表得临界值$U_\alpha = U_{0.05} = 1.645$

根据样本观察值，计算统计观测值

$$u_0 = \frac{\overline{x} - 1000}{\sigma/\sqrt{n}} = \frac{950 - 1000}{100/\sqrt{25}} = -2.5 < -U_{0.05} = -1.645$$

所以统计量落在拒绝域内，即有 95%的把握认为这批产品是不合格的。

注：注意题目是单侧检验还是双侧检验。

3.8.5 正常人的脉搏平均为 72 次/分，现某医生测得 10 例慢性病患者的脉搏（次/分）如下：

54　67　68　78　70　66　67　70　65　69

问在 $\alpha = 0.05$ 水平下，次病患者和正常人的脉搏有无显著差异（已知患者的脉搏服从正态分布）。

解　设患者脉搏数$X \sim N(\mu, \sigma^2)$，$n = 10$.

根据题意，假设 $H_0: \mu = 72 \quad H_1: \mu \neq 72$

因为总体方差未知，选取统计量$T = \frac{\overline{X}-\mu}{S/\sqrt{n}}$

当H_0成立时，$T = \frac{\overline{X}-72}{S/\sqrt{n}} \sim t(n-1)$

对于给定的显著性水平$\alpha = 0.05$，所作检验的模型为模型（Ⅰ），

拒绝域为$\{|t_0| \geq t_{\frac{\alpha}{2}}(n-1)\}$，

即 $P(|T| \geq t_{\frac{\alpha}{2}}(n-1)) = \alpha = 0.05$

所以 $P(T \geq t_{\frac{\alpha}{2}}(n-1)) = \frac{\alpha}{2} = 0.025$

查表得临界值$t_{\frac{\alpha}{2}}(n-1) = t_{0.025}(9) = 2.2622$

根据样本观测值，计算统计量观测值，$\bar{x} = 67.4, s = 5.929$

$$|t_0| = |\frac{\bar{x}-72}{s/\sqrt{n}}| = |\frac{67.4-72}{5.929/\sqrt{10}}| = 2.45 > t_{0.025}(9) = 2.2622$$

所以统计量落在拒绝域内，即有 95%的把握认为病患者和正常人的脉搏有显著差异。

注：过程正确率比较高，算错数的多。

3.8.6 检测某种型号玻璃纸的横向延伸率，测得 25 个数据，经计算得出$\overline{x} = 44.2(\%)$，样本标准差 $s = 4.56(\%)$，假设总体服从正态分布，在显著性 $\alpha = 0.05$ 下，检验玻璃纸的横向延伸率的均值是否为 55%？

解 设玻璃纸的横向延伸率$X \sim N(\mu, \sigma^2)$， $n = 25$.

根据题意，作假设 $H_0: \mu = 55 \quad H_1: \mu \neq 55$

因为总体方差未知，选取统计量$T = \frac{\overline{X}-\mu}{S/\sqrt{n}}$

当H_0成立时，$T = \frac{\overline{X}-55}{S/\sqrt{n}} \sim t(n-1)$

对于给定的显著性水平$\alpha = 0.05$，所作检验的模型为模型（Ⅰ），

拒绝域为 $\{|t_0| \geq t_{\frac{\alpha}{2}}(n-1)\}$

即 $P(|T| \geq t_{\frac{\alpha}{2}}(n-1)) = \alpha = 0.05$

所以 $P(T \geq t_{\frac{\alpha}{2}}(n-1)) = \frac{\alpha}{2} = 0.025$

查表得临界值$t_{\frac{\alpha}{2}}(n-1)=t_{0.025}(24)=2.0639$

根据样本观测值，计算统计量观测值，$\overline{x}=44.2, s=4.56$

$$|t_0|=|\frac{\overline{x}-55}{\frac{s}{\sqrt{n}}}|=|\frac{44.2-55}{\frac{4.56}{\sqrt{25}}}|=11.84>t_{0.025}(24)=2.0639$$

所以统计量落在拒绝域内，即有 95%的把握认为玻璃纸的横向延伸率的均值不为 55%。

注：同上。

3.8.8 某机床原来加工的轴椭圆度的平均值为 0.085mm，标准差为 0.025mm，现在进行技术改革后，在产品中选取 200 个轴进行检验，测得轴椭圆度的平均值为 0.083mm，若假定标准差不变，试在 $\alpha=0.05$ 下检验该厂生产是否正常。

解　设轴椭圆度 X（非正态分布），$E(X)=0.085$，$D(X)=0.025^2$，

$\overline{x}=0.083, n=200.$

根据题意，作假设 $H_0\!:\mu=0.085$　$H_1\!:\mu\neq 0.085$

因为总体方差已知，由中心极限定理，选取统计量$U=\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$

当H_0成立时，$U=\frac{\overline{X}-0.085}{\sigma/\sqrt{n}}\sim N(0,1)$　（近似地）

对于给定的显著性水平$\alpha=0.05$，所作检验的模型为模型（Ⅰ），拒绝域为$\{|u_0|\geq U_{\frac{\alpha}{2}}\}$，即$P(-U_{\frac{\alpha}{2}}<U<U_{\frac{\alpha}{2}})=\Phi(U_{\frac{\alpha}{2}})-\Phi(-U_{\frac{\alpha}{2}})=1-\alpha=0.95$

所以　$\Phi(U_{\frac{\alpha}{2}})=1-\frac{\alpha}{2}=0.975$

查表得临界值$U_{\frac{\alpha}{2}}=U_{0.025}=1.96$

根据样本观察值，计算统计观测值

$$|u_0|=|\frac{\overline{x}-0.085}{\sigma/\sqrt{n}}|=|\frac{0.083-0.085}{0.025/\sqrt{200}}|=1.13<U_{0.025}=1.96$$

所以统计量没有落在拒绝域内，即有 95%的把握认为现在该厂生产是正常的。

注：同上。

3.8.11 为了比较两种枪弹的速度（单位：米/秒），在相同的条件下进行速度测定，算得样本平均值和样本方差为分别为：

枪弹甲：$\mathrm{n_1}=110, \overline{\mathrm{x}}=2805, \mathrm{s_1^2}=120.41$

枪弹乙：$n_2=100, \overline{y}=2680, s_2^2=100.00$.

在显著性水平 $\alpha=0.05$ 下是否可以判定枪弹甲的平均速度比枪弹乙得速度快？

解　设枪弹甲的速度 X（非正态），$E(X)=\mu_1$，$D(X)=\sigma_1^2$，$n_1=110$，

设枪弹乙的速度 Y（非正态），$E(Y)=\mu_2$，$D(Y)=\sigma_2^2$，$n_2=100$，

根据题意做出如下假设检验

$H_0:\mu_1-\mu_2\le 0$　$H_1:\mu_1-\mu_2>0$　与模型　$H_0:\mu_1-\mu_2=0$　$H_1:\mu_1-\mu_2>0$

有相同的拒绝域

由于两个总体方差未知，且样本足够大，

选取统计量$Z=\dfrac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}}$

当H_0成立时，$Z=\dfrac{\overline{X}-\overline{Y}}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}}\sim N(0,1)$（近似地）

对于给定的显著性水平$\alpha=0.05$，所作检验的模型为模型（Ⅱ），拒绝域是$\{z_0\ge U_\alpha\}$，即

$$P(Z>U_\alpha)=1-P(Z\le U_\alpha)=\alpha=0.05$$

所以　$\Phi(U_\alpha)=1-\alpha=0.95$

查表得临界值　$U_\alpha=U_{0.05}=1.645$

根据样本观测值计算，$\overline{x}=2805, s_1^2=120.41, \overline{y}=2680, s_2^2=100.00$,

计算统计量观测值

$$z_0=\frac{\overline{x}-\overline{y}}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}=\frac{2805-2680}{\sqrt{\frac{120.41}{110}+\frac{100}{100}}}=86.37>U_{0.05}=1.645$$

所以统计量落在拒绝域内，即有 95%的把握可以判定枪弹甲得品军速度比枪弹乙得速度快。

注：该题所用检验模型为模型（Ⅱ），注意回答书写规范。

3.8.12 某学校从两个班得学生中随机选取样本，研究完成某课后作业需要的平均

时间，甲班抽 60 名学生，其平均时间$\bar{x}=26.1$，样本方差$s_1^2=144$；乙班抽了 80 名学生，其平均时间$\bar{y}=17.6$，样本方差$s_2^2=110$。设$\alpha=0.05$，能否说明乙班完成时间比甲班完成时间短（假定完成课后作业时间服从正态分布）？

解　甲班学生完成课后作业所需时间 X，$X\sim N(\mu_1,\sigma_1^2)$，

乙班学生完成课后作业所需时间 Y，$Y\sim N(\mu_2,\sigma_2^2)$，

根据题意做出如下假设检验

$H_0:\mu_1-\mu_2\le 0$ $H_1:\mu_1-\mu_2>0$ 与模型 $H_0:\mu_1-\mu_2=0$ $H_1:\mu_1-\mu_2>0$ 有相同的拒绝域。

由于两个总体方差未知，且$n_1=60,n_2=80$，

选取统计量$Z=\frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}}$

当H_0成立时，$Z=\frac{\bar{X}-\bar{Y}}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}}\sim N(0,1)$（近似地）

对于给定的显著性水平$\alpha=0.05$，所作检验的模型为模型（II），拒绝域是$\{z_0\ge U_\alpha\}$，即

$$P(Z>U_\alpha)=1-P(Z\le U_\alpha)=\alpha=0.05$$

所以$\Phi(U_\alpha)=1-\alpha=0.95$

查表得临界值$U_\alpha=U_{0.05}=1.645$

根据样本观测值计算，$\bar{x}=26.1,s_1^2=144,\bar{y}=17.6,s_2^2=110,$

计算统计量观测值

$$z_0=\frac{\bar{x}-\bar{y}}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}=\frac{26.1-17.6}{\sqrt{\frac{144}{60}+\frac{110}{80}}}=4.375>U_{0.05}=1.645$$

所以统计量落在拒绝域内，即有 95%的把握认为乙班完成时间比甲班短。

注：同上，注意计算。

3.8.13 由居民 I 区抽 100 户组成样本，其在目前得房子中平均居住得时间是$\bar{x}=$

35 个月，样本方差$s_1^2=900$；在居民 II 区抽 120 户组成样本，其在目前得房子中平均居住得时间是$\overline{y}=49$ 个月，样本方差$s_2^2=1000$.在 $\alpha=0.05$ 水平下，是否可以判定 I 区居民平均居住时间比 II 区短？

解　设在居民 I 区居住时间 X（非正态），$E(X)=\mu_1, D(X)=\sigma_1^2, n_1=100$，

设在居民 II 区居住时间 Y（非正态），$E(Y)=\mu_2, D(Y)=\sigma_2^2, n_2=120$，

根据题意做出如下假设检验

$H_0:\mu_1-\mu_2\geq 0$ $H_1:\mu_1-\mu_2<0$ 与模型 $H_0:\mu_1-\mu_2=0$ $H_1:\mu_1-\mu_2<0$ 有相同的拒绝域。

由于两个总体方差未知，且样本足够大，

选取统计量$Z=\dfrac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}}$

当H_0成立时，$Z=\dfrac{\overline{X}-\overline{Y}}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}}\sim N(0,1)$（近似地）

对于给定的显著性水平$\alpha=0.05$，由于所作检验的模型为模型（III），拒绝域是$\{z_0\leq -U_\alpha\}$，即

$$P(Z<-U_\alpha)=\Phi(-U_\alpha)=1-\Phi(U_\alpha)=0.05$$

所以$\Phi(U_\alpha)=1-\alpha=0.95$

查表得临界值$U_\alpha=U_{0.05}=1.645$

根据样本观测值计算，$\overline{x}=35, s_1^2=900, \overline{y}=49, s_2^2=1000$，

计算统计量观测值

$$z_0=\frac{\overline{x}-\overline{y}}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}=\frac{35-49}{\sqrt{\frac{900}{100}+\frac{1000}{120}}}=-3.36<-U_{0.05}=-1.645$$

所以统计量落在拒绝域内，即有 95%的把握认为可以判定 I 区居民平均居住时间比 II 区短。

注：这道题算错数的同学比较多。

3.9 线性回归分析 参考答案

3.9.3 有人认为，企业利润水平和它的研究费用之间存在近似地线性关系，下表所列资料能否证实这种判断(α = 0.05).

研究费 x_i (万元)	10	10	8	8	8	12	12	12	11	11
利润 y_i (万元)	100	150	200	180	250	300	280	310	320	300

解　因为$\sum_{i=1}^{10} x_i = 102, \sum_{i=1}^{10} y_i = 2390$，所以$\bar{x} = 10.2, \bar{y} = 239$，

又 $\sum_{i=1}^{10} x_i^2 = 1066, \sum_{i=1}^{10} y_i^2 = 624300, \sum_{i=1}^{10} x_i y_i = 25040$，所以

$$L_{xx} = \sum_{i=1}^{10} x_i^2 - \frac{1}{10}\left(\sum_{i=1}^{10} x_i\right)^2 = 1066 - \frac{1}{10} \times 102^2 = 25.6$$

$$L_{xy} = \sum_{i=1}^{10} x_i y_i - \frac{1}{10}\left(\sum_{i=1}^{10} x_i\right)\left(\sum_{i=1}^{10} y_i\right) = 25040 - \frac{1}{10} \times 102 \times 2390 = 662$$

$$L_{yy} = \sum_{i=1}^{10} y_i^2 - \frac{1}{10}\left(\sum_{i=1}^{10} y_i\right)^2 = 624300 - \frac{1}{10} \times 2390^2 = 53090$$

所以，$\hat{b} = \frac{L_{xy}}{L_{xx}} = 25.86, \hat{a} = \bar{y} - \hat{b}\bar{x} =- 24.77$，

根据题意，要检验的假设为$H_0: b = 0 \quad H_1: b \neq 0$

又　$n = 10, \alpha = 0.05$

查相关系数显著性检验表得：$r_\alpha(n-2) = r_{0.05}(8) = 0.632$，

而 $|r_0| = |\frac{L_{xy}}{\sqrt{L_{xx}L_{yy}}}| = |\frac{662}{\sqrt{25.6 \times 53090}}| = 0.5678 < 0.632$，

接受H_0，即 x 与 y 之间的线性关系不显著，所以不能确定企业利润水平和它的研究经费之间存在显著线性关系。

注：这种题要算的数比较多，看准了数再算，算错数的同学非常多！

3.9.7 为了研究小麦某种害虫的生长规律，测得害虫从产卵到孵化成幼虫的天数 N 以及孵化期内每日平均温度的算术平均数 T（如下表）

编号	1	2	3	4	5	6	7	8
平均气温 T_i	11.8	14.7	15.4	16.5	17.1	18.3	19.8	20.3
孵化期 N_i	30.4	15.0	13.8	12.7	10.7	7.5	6.8	5.7

试求 N 对 T 的回归方程，并进行显著性检验(α = 0.05).

解　因为$\sum_{i=1}^{8} T_i = 133.9$，$\sum_{i=1}^{8} N_i = 102.6$，所以$\overline{T} = 16.7375, \overline{N} = 12.825$，

又 $\sum_{i=1}^{8} T_i^2 = 2296.17$, $\sum_{i=1}^{8} N_i^2 = 1750.36$, $\sum_{i=1}^{8} T_i N_i = 1571.86$

所以

$$L_{TT} = \sum_{i=1}^{8} T_i^2 - \frac{1}{8}(\sum_{i=1}^{8} T_i)^2 = 2296.17 - \frac{1}{8} \times 133.9^2 = 55.02$$

$$L_{TN} = \sum_{i=1}^{8} T_i N_i - \frac{1}{8}(\sum_{i=1}^{8} T_i)(\sum_{i=1}^{8} N_i) = 1571.86 - \frac{1}{8} \times 133.9 \times 102.6 =- 145.42$$

$$L_{NN} = \sum_{i=1}^{8} N_i^2 - \frac{1}{8}(\sum_{i=1}^{8} N_i)^2 = 1750.36 - \frac{1}{8} \times 102.6^2 = 434.52$$

所以　$\hat{b} = \frac{L_{TN}}{L_{TT}} =- 2.643, \hat{a} = \overline{N} - \hat{b}\overline{T} = 57.06$

根据题意，要检验的假设为$H_0: b = 0 \quad H_1: b \neq 0$

又 $n = 8, \alpha = 0.05$

查相关系数显著性检验表得：$r_\alpha(n-2) = r_{0.05}(6) = 0.707$

而 $|r_0| = |\frac{L_{TN}}{\sqrt{L_{TT}L_{NN}}}| = |\frac{145.42}{\sqrt{55.02 \times 434.52}}| = 0.9405 > 0.707$

拒绝H_0，即 T 与 N 之间的线性关系显著，N 对 T 的线性回归方程为：

$$\hat{N} = 57.06 - 2.643T$$

注：线性回归方程是 N 对 T，有的同学写着写着就变 X 了。

3.9.9 下表是 6 中不同的钢标本资料.

强力 x	**1**	**2**	**3**	**4**	**5**	**6**
伸长率 y	**15**	**35**	**41**	**63**	**77**	**84**

试求（1）y 对 x 线性回归方程；（2）检验线性回归方程的显著性($\alpha = 0.05$)。

解　因为　$\sum_{i=1}^{6} x_i = 21, \sum_{i=1}^{6} y_i = 315$，所以$\overline{x} = 3.5, \overline{y} = 52.5$，

又　$\sum_{i=1}^{6} x_i^2 = 91$, $\sum_{i=1}^{6} y_i^2 = 20085, \sum_{i=1}^{6} x_i y_i = 1349$，

所以

$$L_{xx}=\sum_{i=1}^{6}x_i^2-\frac{1}{6}\left(\sum_{i=1}^{6}x_i\right)^2=91-\frac{1}{6}\times 21^2=17.5$$

$$L_{xy}=\sum_{i=1}^{6}x_iy_i-\frac{1}{6}\left(\sum_{i=1}^{6}x_i\right)\left(\sum_{i=1}^{6}y_i\right)=1349-\frac{1}{6}\times 21\times 315=246.5$$

$$L_{yy}=\sum_{i=1}^{6}y_i^2-\frac{1}{6}\left(\sum_{i=1}^{6}y_i\right)^2=20085-\frac{1}{6}\times 315^2=3547.5$$

所以 $\hat{b}=\frac{L_{xy}}{L_{xx}}=14.086$, $\hat{a}=\overline{y}-\hat{b}\overline{x}=3.2$，

根据题意，要检验的假设为 $H_0: b=0\ \ H_1: b\neq 0$

又 $n=6, \alpha=0.05$

查相关系数显著性检验表得：$r_\alpha(n-2)=r_{0.05}(4)=0.811$

而 $|r_0|=|\frac{L_{xy}}{\sqrt{L_{xx}L_{yy}}}|=|\frac{246.5}{\sqrt{17.5\times 3547.5}}|=0.9893>0.811$

拒绝H_0，即 x 与 y 之间线性关系显著，因此 x 与 y 之间的线性回归方程为：

$$\hat{y}=3.2+14.086x$$

注：解题过程并不复杂，计算稍微费力，算错的同学不少，同学们平时要看仔细、多练习自己的计算能力哈！