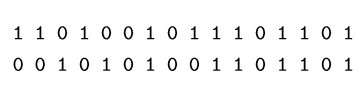
实验三 R语言与统计推断

【实验目的】

1. 掌握用R语言进行点估计的矩法估计和极大似然估计。
2. 掌握用R语言进行的单正态总体的均值和方差的估计、两正态总体的参数估计、比率的区间估计。
3. 掌握用R语言实现单个和两个正态总体的参数和非参数假设检验。

【实验内容与实现】

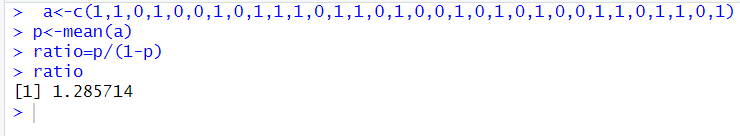
1. 对某个篮球运动员记录其在一次比赛中投篮命中与否, 观测数据如下：



请利用矩法估计求这位篮球运动员投篮的成败比，用R实现**。**

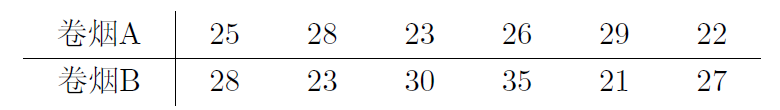
由题目可知，数据服从一个二项分布。

假设投篮成功的概率为P，那么失败的概率为1-P，投篮的成败比为P/1-P。利用观测数据对P进行矩法估计如下：



利用一阶原点矩估计出的结果如图所示。成败比为1.285714

1. 某卷烟厂生产两种卷烟A和B，现分别对两种香烟的尼古丁含量进行6次试验，结果如下

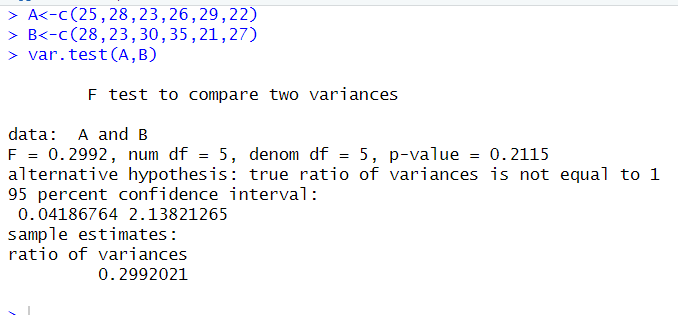


若香烟的尼古丁含量服从正态分布，

1) 问两种卷烟中尼古丁含量的方差是否相等?

原假设，备择假设，统计量，检验方法

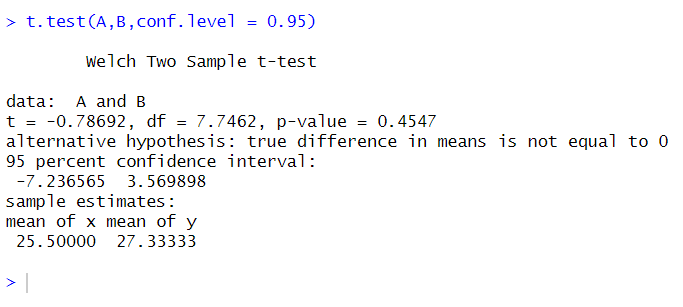
由题可知，两组观测数据相互独立，使用F检验对这两个总体的方差是否相等进行检验。假设H0：两种卷烟的方差相等；H1：两种卷烟的方差不等



结论：三句话，结合题目

图中P-value为0.2115，表示不能拒绝原假设，因此两种卷烟的方差相等。

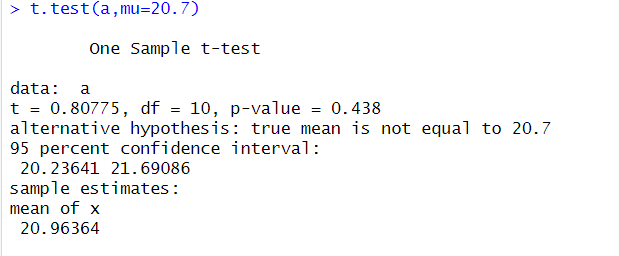
2) 试求两种香烟的尼古丁平均含量差的95%置信区间。



该结果的备择假设是平均含量差不等于0，由于P-value为0.45，因此无法拒绝原假设，两种卷烟的平均含量差为0。在这个情况下的95%的置区间为-7.2到3.57

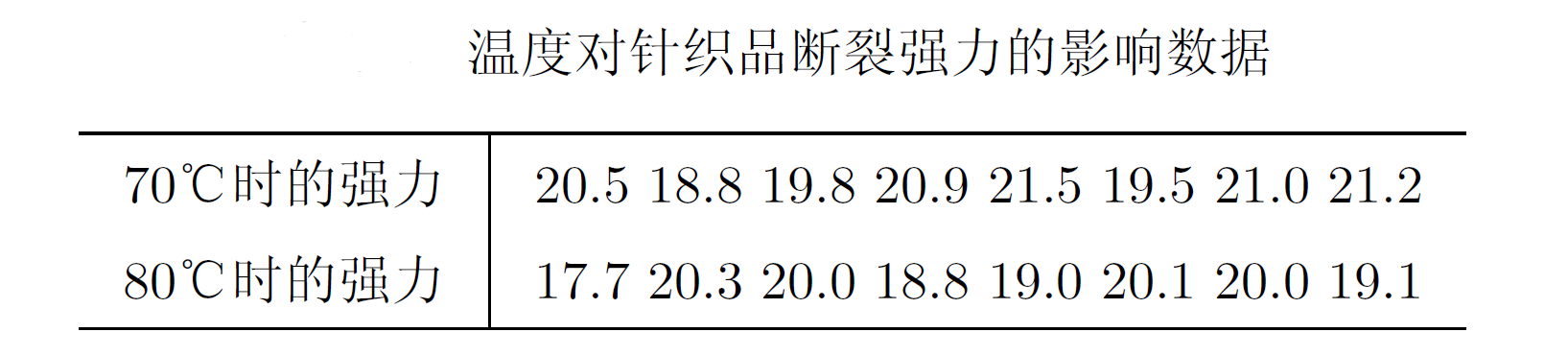
1. 已知某种水样中CaCO3的真值为20.7mg/L，现用某种方法重复测定该水样11次，CaCO3的含量为: 20.9, 20.41, 20.10, 20.00, 20.19, 22.60, 20.99, 20.41, 20, 23, 22。假设水样中CaCO3服从正态分布， 问用该法测定的CaCO3含量是否可信? (显著性水平为0.05)

由题可知，已知u=20.7，对该样本进行t检验即可测定总体均值为20.7是否显著。假设：H0：测得CaCO3含量可信；H1：测得CaCO3含量不可信。



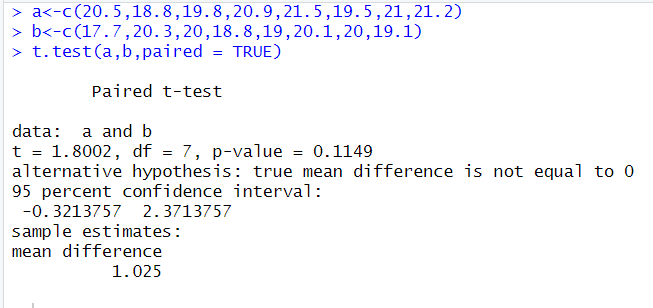
由图可知，P-value为0.438大于0.05，故无法拒绝原假设，因此测得的CaCO3的含量可信。

1. 在针织品漂白工艺过程中，要考虑温度对针织品断裂力（主要质量指标）的影响，为了比较70℃与80℃的影响有无差别，在这两个温度下，分别用相同的针织品重复做了8次试验，得数据如表所示（单位：N）：根据经验，温度对针织品断裂强度的波动没有影响。



问在70℃时的平均断裂力与80℃时的平均断裂力之间是否有显著差别？假定断裂强力服从正态分布

由题可知，温度在70摄氏度和80摄氏度下对丝织品断裂强度得波动不是相互独立的。二者的实际差别即为总体均值的差别。那么进行以下假设。H0：70℃与80℃的影响无差别 ；H1：70℃与80℃的影响有差别



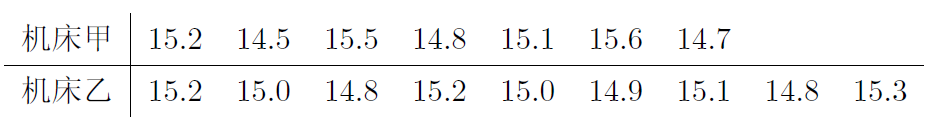
由图可知，P-value=0.1149,>0.05因此无法拒绝原假设，那么70℃与80℃的影响无差别。

f

【作业】

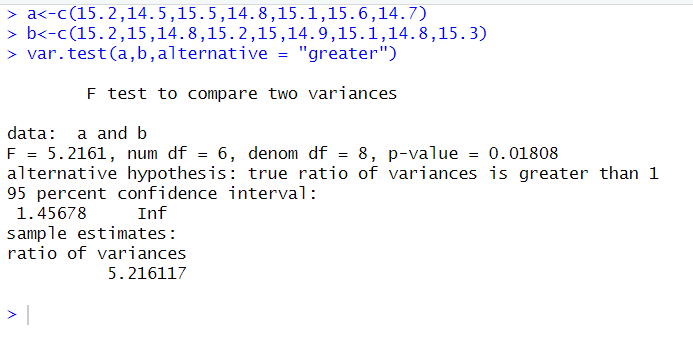
要求：用R语言求解下列问题，并给出完整的解题步骤。

1. 有两台机床生产同一型号的滚珠，根据以往经验知，这两台机床生产的滚珠直径都服从正态分布. 现分别从这两台机床生产的滚珠中随机地抽取7个和9个，测得它们的直径如下(单位: mmm)



试问机床乙生产的滚珠的方差是否比机床甲生产的滚珠直径的方差小?

由题可知，两机床生产滚珠直径应当是相互独立的，要比较二者方差的大小进行F检验即可。假设：H0：机床乙生产的滚珠方差比机床甲大；H1：机床乙生产的滚珠方差比机床甲大



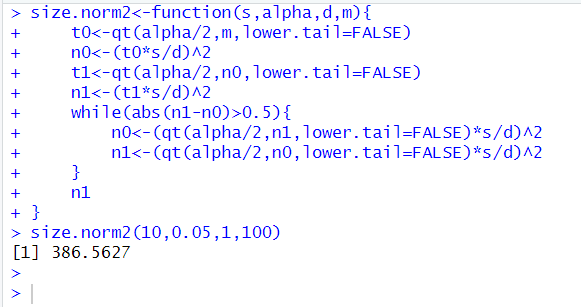
由图中P-value为0.01808小于0.05，因此拒绝原假设，那么机床乙生产的滚珠方差比机床甲大。

1. 某公司生产了一批新产品, 产品总体服从正态分布, 现要估计这批产品的平均重量, 最大允许误差为1, 样本标准差为10, 试问在0.95的置信度下至少要抽取多少个产品?

由于总体的方差未知，因此样本最终符合的是一个t分布，根据t的表达式

用t把n表示出来n=[t\*s/(x-u)]^2，先计算t分布下的置信区间，再对n进行计算即可。

那么用R语言实现如下：

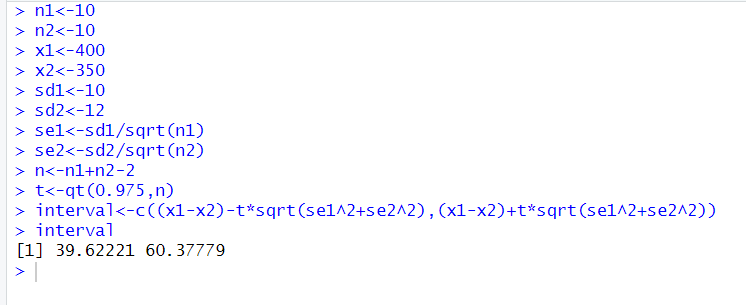


至少要抽取387个样本。

1. 某饲养场分别使用传统和新饲料喂养牛，假设牛的体重服从正态分布。一定时间后随机取10头样本。传统饲料样本平均400KG，标准差10KG；新饲料平均350KG，标准差12KG。问置信度为95%的两种饲料喂养的牛体重差异置信度区间是多少？

首先计算两个样本的标准误差SE，然后利用

置信区间=（X1-X2）+/- t\*SE计算出0.95水平下的置信区间即可。



最终的0.95置信度下的体重差异的置信区间为39.62221到60.37779。