**1、对以往数据分析结果表明，当机器调整得良好时，产品的合格率为90%，而当机器发生某一故障时，其合格率为30%。每天早上机器开动时，机器调整良好的概率为75%，试求已知某日早上第一件产品是合格品时，机器调整良好的概率是多少？(10')**

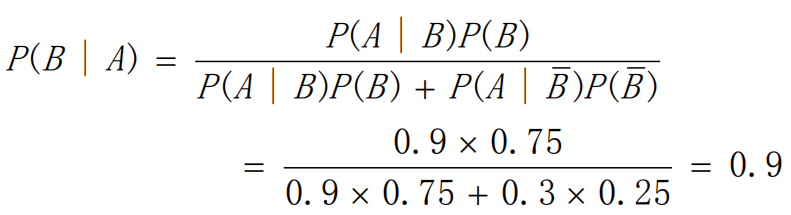
答案：

解：设 A 为事件“产品合格”，B 为事件“机器调整良好”（2’）

已知:P(A|B)=0.9，P(A|B)=0.3,P（B）=0.75，（2’）

（以上两步只要把事情说清楚即可，可以不用 A、B 进行事件代称）

所需求的概率为 P(B|A)，由贝叶斯公式：



（公式 5’，答案 1’）

**2、某零件用两种工艺加工，第一种工艺有三道工序，各道工序出现不合格品的概率分别是 0.3,0.2,0.1；第二种 工艺有两道工序，各道工序出现不合格品的概率分别为0.3,0.2，试问：**

**（1）用哪种工艺加工得到合格品的概率较大些？(10')**

**（2）第二道工艺两道工序出现不合格品的概率都是0.3时，情况又如何？(10')**

答案：

1. 记事件Ai为用第i中工艺得到的合格品，i=1,2，···

由于各道工序可看做是独立工作的，所以：

P（A1）=0.7 X 0.8 X 0.9=0.504

P（A2）=0.7 X 0.8=0.56

所以第二种工艺得到合格品的概率大些。

1. 当第二道工艺的两道工序出现不合格的概率都是0.3时

P（A2）=0.7 X 0.7=0.49

所以第一种工艺得到合格品的概率大些。

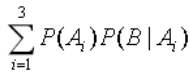
**3、雨伞掉了。落在图书馆的概率为50%，这种情况下找回的概率为0.80，落在教室的概率为30%，这种情况下找回的概率为0.60；落在商场的概率为20%，这种情况下找回的概率为0.05，求找回雨伞的概率。(10')**

答案：

设找回为事件B，A1,A2,A3分别为雨伞落在图书馆，教室，商场。

所以B=

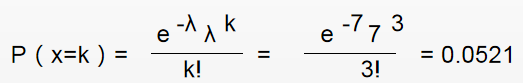
因为P（A1）=0.05，P（A2）=0.3，P（A3）=0.2，P（B|A1）=0.8，P（B|A2）=0.6，P（B|A3）=0.05

所以P（B）==0.5 X 0.8 + 0.3 X 0.6 + 0.2 X 0.05=0.59

**4、一个交通路口的一个月交通事故次数服从平均值为7的Poisson distribution。下一个月观察到3次交通事故的概率为多少？(12')**

答案：

解：根据泊松分布 ，k=3



**5、研究人员对锌水平是否会因饮食中补充钙而发生改变感兴趣。研究人员用20只小鼠做实验（假设已知饮食中没有补充钙小鼠血液锌的浓度是 1 mg/ml），在它们的饮食中补充钙，并测量血液中锌的浓度（mg/ml）分别是：1.31, 1.45, 1.12, 1.16, 1.30, 1.50, 1.20, 1.22, 1.42, 1.14, 1.23, 1.59, 1.11, 1.10, 1.53, 1.52, 1.17, 1.49, 1.62, 1.29 。**

**（1）计算平均值和样本标准差，并计算平均值的标准误差？(8')**

**（2）95%的置信区间是多少？(8')**

**（3）小鼠饮食中补充钙是否会改变它血液中锌的浓度，计算P值（0.05作为阈值）? (8')**

答案：

（1）1.32，0.17，0.04 （2’,3’,3’）

（2）[1.24, 1.4] （8’）

（3）P值为0.005 。小鼠饮食中补充钙会改变它血液中锌的浓度。（8’）

代码：

> zinc <- c(1.31, 1.45, 1.12, 1.16, 1.30, 1.50, 1.20, 1.22, 1.42, 1.14, 1.23, 1.59, 1.11, 1.10, 1.53, 1.52, 1.17, 1.49, 1.62, 1.29)

**> #Answer 1**

> zinc\_m<-mean(zinc);zinc\_m

[1] 1.3235

> zinc\_sd<-sd(zinc);zinc\_sd

[1] 0.174997

> zinc\_se<- zinc\_sd/sqrt(length(zinc));zinc\_se

[1] 0.03913052

**> #Answer 2**

> c(zinc\_m-qt(0.975,length(zinc)-1)\*zinc\_se,zinc\_m+qt(0.975,length(zinc)-1)\*zinc\_se)

1. 1.241599 1.405401

**> #Answer 3**

> t.test(zinc,mu = 1, alternative = "two.sided",conf.level = 0.95)

One Sample t-test

data: x

t = 8.2672, df = 19, p-value = 1.025e-07

alternative hypothesis: true mean is not equal to 1

95 percent confidence interval:

1.241599 1.405401

sample estimates:

mean of x

1.3235

**6、已知某校2万人体检测得所有人的平均体重为59kg，标准差为25kg，从中抽取20人的样本体重值如下表所示：**

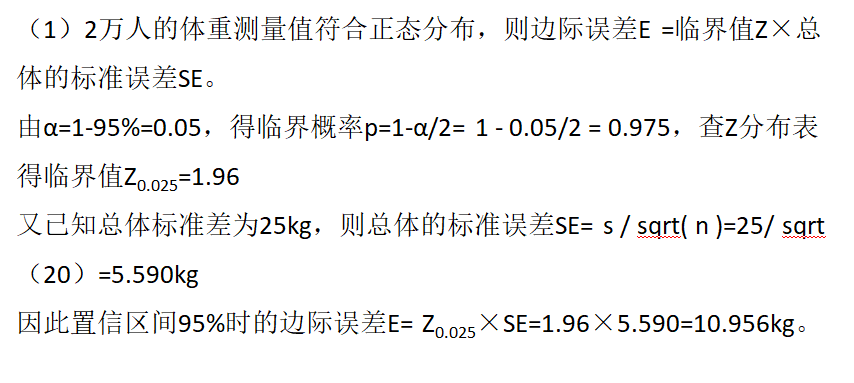
|  |  |
| --- | --- |
| **样本编号** | **体重（kg）** |
| **1** | **61.5** |
| **2** | **54** |
| **3** | **60.5** |
| **4** | **54.5** |
| **5** | **52.5** |
| **6** | **67.5** |
| **7** | **63** |
| **8** | **57.5** |
| **9** | **65.5** |
| **10** | **58** |
| **11** | **60** |
| **12** | **70** |
| **13** | **53** |
| **14** | **64** |
| **15** | **56** |
| **16** | **66** |
| **17** | **70.5** |
| **18** | **57.5** |
| **19** | **49** |
| **20** | **59.5** |

**（1）求置信区间95%时的边际误差E？（8’）**

**（2）求样本均值落在50~63kg之间的概率？（8’）**

**（3）求样本均值95%的置信区间？（8’）**

答案：



（2）由中心极限定理可得，, 所以

R code:

> stdvar <- 25/sqrt(20)

> pnorm(63, mean=59, sd=stdvar)

[1] 0.7628628

> pnorm(50, mean=59, sd=stdvar)

[1] 0.05370232

> pnorm(63, mean=59, sd=stdvar)-pnorm(50, mean=59, sd=stdvar)

[1] 0.7091605

所以，样本均值落在[50,63]之间的概率为0.709。

（3）=0.05，/2=0.025, 1-=0.975,由样本均值估值总体均值时，95%置信区间为：

95% CI：

R code:

> hospital<-c(61.5,54,60.5,54.5,52.5,67.5,63,57.5,65.5,58,60,70,53,64,56,66,70.5,57.5,49,

59.5)

> hospital\_mean <- mean(hospital)

> hospital\_sd <- sd(hospital)

> a <- hospital\_mean - qnorm(0.975)\*stdvar

> a

[1] 49.04347

> b <- hospital\_mean + qnorm(0.975)\*stdvar

> b

[1] 70.95653

所以，总体均值95%的置信区间为 [49.04，70.96]。