Homework #5

1. **(12 points)** All cigarettes presently on the market have an average nicotine content of 1.6 mg per cigarette. A company that produces cigarettes want to test if the average nicotine content of a cigarette is 1.6 mg. To test this, a sample of 36 of the company's cigarettes were analyzed.

   a) If it is known that the standard deviation of a cigarette's nicotine content is 0.3 mg, what conclusions can be drawn, at the 1 percent level of significance, if the average nicotine content of the 36 cigarettes is 1.475?

   H0: u=1.6
   H1:u!=1.6
   P=2(1-Phi(z))=2(1-Phi(z=x-u/(s/sqrt(n))))
   =2(1-Phi(0.125*6/0.3))=2(1-Phi(2.5))=2*0.0062=0.0124
   accept H0

   z=x-u/(s/sqrt(n))=-2.5
   reject region is [-Z0.005, Z0.005]=[-2.58, 2.58]

   b) What is the P-value for the hypothesis test in (a)?

   P=2(1-Phi(z))=2(1-Phi(z=x-u/(s/sqrt(n))))
   =2(1-Phi(0.125*6/0.3))=2(1-Phi(2.5))=2*0.0062=0.0124

2. **(12 points)** In a sample of 500 users of a computer program, 162 said they are satisfied. Construct a 95% confidence interval for the population proportion.

   z-static p=162/500=0.324, s=sqrt(p(1-p))=0.468
   z=P-p/(sqrt(p(1-p))/n)
   P(p-Z0.025s/sqrt(n)<z<p+Z0.025s/sqrt(n))=0.95
   [0.324-1.96*0.021, 0.324+1.96*0.021]=[0.283, 0.365]

3. **(12 points)** The table below shows the number of students absent from school on particular days in the week.

Day	**M**	**Tu**	**W**	**Th**	**F**
Number	125	88	85	94	108

   a) Find the expected frequencies (expected numbers of absent students) if it is assumed that students are equally likely to be absent on each working day.

   iid

   E(x)=sum x/n=100

b) Test, at the significance level of 5%, the null hypothesis that students are equally likely to be absent on each working day. Use chi-squared goodness of fit test.

H0: uniform distribution
H1: not uniform
E=100 for each day
chi^2=sum (O-E)^2/E=25^2+12^2+15^2+6^2+8^2/100=10.94
P-value=1-chi2cdf(chi^2,k-1=4)=0.05 -> chi^2,0.05,4=9.49
10.94>9.49, reject H0

4. **(10 points)** Cancer researchers have tested the effectiveness of a new drug on rats. A group of rats with tumors were given the drug and initially their tumors shrank to be undetectable. Listed below are the times (in days) until cancer in rats developed resistance to the drug so that tumors reappeared.

101, 104, 77, 104, 96, 82, 70, 89, 91, 103, 93, 85, 104, 104, 81, 67, 104, 87, 104, 89, 78, 104, 86, 76, 103, 102, 80, 45, 94, 104, 104, 76, 80, 72, 73

Use chi-squared distribution to calculate the 95% confidence interval on the population standard deviation of these times of tumor reappearance.

u=sum x/n=88.9
s=14.2
sqrt((n-1)*s^2/chi^2_0.975,35-1)<sig<sqrt((n-1)*s^2/chi^2_0.025,35-1)

5. **(10 points)** Three identical six-sided dice, each with faces marked 1 to 6, are rolled 100 times. At each rolling, a record is made of the number of dice whose score is 5 or 6. The results are as follows.

# of dice with score 5 or 6	0	1	2	3
# of rolls	31	41	19	9

Please explain whether these results are consistent with the **unbiased dice** hypothesis. Use chi-squared goodness of fit test with α=3% to justify your conclusions.

binomial, P(c=0)=1*(2/3)^3=0.296,
P(c=1)=C1,3 (2/3)^2*1/3=0.444
P(c=2)=C2,3*(2/3)(1/3)^2=0.222
P(c=3)=(1/3)^3=0.037

E(x=0)=100*P(c=0)=29.6=30
E(x=1)=100*P(c=1)=44.4=44
E(x=2)=100*P(c=2)=22.2=22
E(x=3)=100*P(c=3)=3.7=4

H0: unbiased dice-> uniform dist
chi^2=sum(O-E)^2/E=6.897
chi^_0.03,99=8.95>6.897, accept H0

6. **(14 points)** A polymer is manufactured in a batch chemical process and viscosity measurements are normally made on each batch. The sample mean and standard deviation of 50 batch viscosity measurements are 750.2 and 19.13, respectively. A process change is made which involves switching the type of catalyst used in the process. Following the process change, 30 batch viscosity measurements are taken and the sample mean and standard deviation are 756.88 and 21.28, respectively.

a) At the significance level of 10%, can you conclude that there is a significant difference between population means before and after the process was changed?

H0:u0-u1=0
t=X1-X2-delta/sqrt(s1^2/n1+s2^2/n2)=-1.41
CI: [0-t0.05,50+30-2,0+t0.05,50+30-2]=[-1.66, 1.66]

P-value =2(1-Phi(z))=2(1-
P(z<1.41))=2(1-0.921)=0.158>0.1
accept H0

b) Find the P-value of the null hypothesis that process change did not have any impact on the population mean.

P-value =2(1-Phi(z))=2(1-P(z<1.41))=2(1-0.921)=0.158>0.1

use t-distribution
2(1-P(t78<1.41))

c) Find a 90% confidence interval on the difference in mean batch viscosities resulting from the process change.

Z=X1-X2-delta/sqrt(s1^2/n1+s2^2/n2)

P(Z0.05<Z<Z0.95)=0.9
X1-X2-sqrt(s1^2/n1+s2^2/n2)*Z0.05<detla<X1-X2
+sqrt(s1^2/n1+s2^2/n2)*Z0.05
-6.68-4.74*1.65<del<-6.68+4.74*1.65
-14.5<del<1.14

for t-distribution
-6.68-t0.05,80-2*4.74<del<-6.68+t0.05,80-2