Homework #4

1. (10 points) A sample of size 100, which has the sample mean $\bar{X} = 500$, was drawn from a population with an unknown mean μ and the standard deviation $\sigma = 80$.

a) What is the probability that the population mean will be in the interval (480, 510)?

Z=X-u/s/sqrt(n)

P(480<u<510)=P(-20/8<Z<10/8)
=P(-2.5<Z<1.25)=Phi(1.25)-Phi(-2.5)=0.8944-0.0063=0.8881

b) Give the 95% confidence interval for the population mean.

P(X-Z0.025 s/sqrt(n)<u<X+Z0.025s/sqrt(n))=0.95
Z0.025=1.96, [500-1.96*8, 500+1.96*8]

2. (10 points) Find the maximum likelihood estimator for λ in a sample of size *n* drawn the Poisson distribution

$$f(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

Log(L)=sum(xi)log\lambda-\lambda*n - log(xi!...xn!)
dlog L/d\lambda=sum(xi)/\lambda-n=0
\lambda=sum(xi)/n

3. (18 points) Among all the computer chips produced by a certain factory, 6% are defective. A sample of 400 chips is selected for inspection.

a) Use Matlab or other computational software to find the exact probability that this sample contains between 20 and 25 defective chips (including 20 and 25)? Hint: the number of defective chips follows the binomial distribution.

sum binomial (20~25, 400, 0.06)

b) Use the central limit theorem to approximate the same probability. Hint: to account for the rounding error integrate the appropriate normal distribution between 19.5 and 25.5.

u=np=24,s=sqrt(np(1-p))=4.75
p(19.5<x<25.5)=P(-4.5/4.75<Z<1.5/4.75)
=P(-0.95<Z<0.32)=0.6255-0.1711=0.4544

c) Suppose that 40 inspectors independently from each other collected samples of 400 chips each. What is the probability that at least 14 inspectors will find between 20 and 25 defective chips in their samples? Hint: use normal distribution

CDF at 13.5 (instead of 14) to approximate the binomial distribution.

P~binomial(40, 0.454)
u=np=19, s=sqrt(np(1-p))=3.15
P(x>=14)=1-P(x<14)=1-P(z<-5/3.15=-1.59)=0.944

4. (10 points) (10 points) The elasticity of a polymer is affected by the concentration of a reactant. When low concentration is used, the true mean elasticity is 55, and when high concentration is used the mean elasticity is 60. The standard deviation of elasticity is 4 in the first case and 5 in the second case. Random samples of sizes 16 and 9 correspondingly are taken, find the probability that $\bar{X}_{high} - \bar{X}_{low} \geq 4$.

P(Xhigh-Xlow)=P(Z=(Xhigh-Xlow)-(u_high-u_low)/sqrt(s_high^2/n_high
+s_low^2/n_low)=P(Z>4-(60-55)/sqrt(5^2/9+4^2/16)))=P(Z>-0.51)=0.685

5. (12 points) To estimate the copy number of a specific protein, a laboratory has done multiple measurements:2310, 2320, 2010, 10800, 2190, 3360, 5640, 2540, 3360, 11800, 2010, 3430, 10600, 7370, 2160, 3200, 2020, 2850, 3500, 10200, 8550, 9500, 2260, 7730, 2250

a) Find a point estimate of the mean protein copy number.

u=sum x/n=4958.4

b) Find a point estimate of the standard deviation of the protein copy number
s=sqrt(sum (xi-u)^2/n-1)=3420.5

c) What is approximately the standard error of the estimate of the mean protein copy number obtained in part a)

SE=s/sqrt(n)=684

d) Find a point estimate for the proportion of readings that are less than 5000.
16/25

e) Find 95% confidence intervals for the point estimate in part d)

<5000 or > 5000, bernouli (p), u=p=0.64, s=sqrt(p(1-p))=0.48
P(X-Z0.025s/sqrt(n)<z<XZ0.025s/sqrt(n))=0.95
[0.64-1.96*0.48/5, 0.64+1.96*0.48/5]=[0.45,0.83]

f) Use the computer to plot the histogram and the box-and-whisker diagram for the sample

hist(sample)
boxplot(sample)