Stat13 Homework 5

http://www.stat.ucla.edu/~dinov/courses_students.html

(50 points, student scores will be converted to scores out of 100)

Suggested Solutions

HW_5_1 <20 points>

E(x)=-0.165
SD(x)=6.84 or 6.97
Pr=0.411 for 0.431
For n=100
E(x_bar)=-0.165
SD(x_bar)=0.684 or 0.697
Pr=0.405 or 0.406
For n=1000
E(x_bar)=-0.165
SD(x_bar)=0.216 or 0.22
Pr=0.222 or 0.227
For n=5000
E(x_bar)=-0.165
SD(x_bar)=0.097 or 0.099
Pr=0.044 or 0.047
For n=10000
E(x_bar)=-0.165
SD(x_bar)=0.0684 or 0.0697
Pr=0.0079 or 0.009.

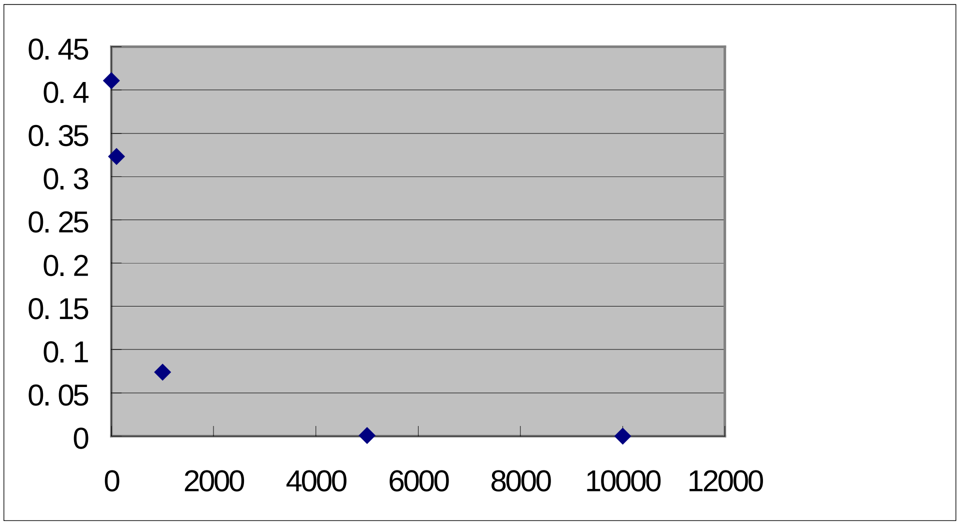
(1) <4 points>
E(X)=-5*0.589+5*0.346+10*0.06+60*0.005=-0.315 <1 point>
SD(X)=sqrt((-5+0.315)^2*0.589+(5+0.315)^2*0.346+(10+0.315)^2*0.06+(60+0.315)^2*0.005)= 6.875738 <2 points>

Prob(positive return in a single bet)=1-0.589=0.411 <1 point>

(2) <4 points>
E(X_bar)=E(X)=-0.315 <1 point>
SE(X_bar)=SD(X)/sqrt(100)=0.688 <1 point>
X_bar is normally distributed, based on Central Limit Theorem. <1 point>
P(X_bar>0)=P(z > (0+0.315)/0.688)=P(z > 0.459)=1-0.6769=0.3231 <1 point>
(3) <9 points>
For 1000 bets:
E(X_bar)=-0.315
SE(X_bar)=0.218

P(X_bar > 0)=P(z >1.44)=0.074 <3 points>
For 5000 bets:
E(X_bar)=-0.315
SE(X_bar)=0.097
P(X_bar > 0)=P(z >3.25)=0.00058 <3 points>
For 10000 bets:
E(X_bar)=-0.315
SE(X_bar)=0.0688
P(X_bar > 0)=P(z >4.58)=0.00000234 <3 points>

(4) <3 points>
<2 points for the scatter plot>
The higher the number of independent bets, the less likely to make a net gain. <1 point>



HW_5_2 <8 points>
(1) <2 points>
SE(X1_bar – X2_bar)=
sqrt(sd^2(X1)/n1+sd^2(X2)/n2)=sqrt(1.91*1.91/40+1.43*1.43/58)=0.356
(2) <2 points>
E(X1_bar – X2_bar)=m1-m2=3.1 <1 point>
Two standard error interval: [3.1-2*0.356, 3.1+2*0.356] = [2.388, 3.812] <1 point>
(3) <4 points>
According to the study:
The difference between the mean of the two groups is normally distributed with mean of 3.1 and standard error of 0.356. The two standard error interval is totally above 0. It is significant that group one (sexual content) has a higher mean of remembering than group two (general audience content). The result of the study can be summarized as: Sexual content ads are more likely to be remembered, at least in the student-aged population.

HW_5_3 <12 points>
(1) <3 points, 1 point each>
one sample:
mean=E(X)=6.75
s.d.=sd(X)=1.1
four sample:
mean=E(X)=6.75
s.d.=sd(X)/2=0.55
16 sample:
mean=E(X)=6.75
s.d.=sd(X)/4=0.275
(2) <2 point>
They differ in standard deviation. Because with larger sample, variance of single units tend to be averaged out. The bigger the sample size, the smaller the standard error.
(3) <7 points>
Y=1.3X-1.3 <1 point>
E(Y)=1.3*E(X)-1.3=7.475 <1 point>
SD(Y)=1.3*SD(X)=1.43 <1 point>
E(X1+X2+X3+Y1+Y2+Y3+Y4+Y5)
=E(X1)+E(X2)+E(X3)+E(Y1)+E(Y2)+E(Y3)+E(Y4)+E(Y5)
=3*6.75+5*7.475
=57.625 <2 points>
SD(X1+X2+X3+Y1+Y2+Y3+Y4+Y5)
=sqrt(sd(X1)+^2+sd(X2)^2+sd(X3)^2+sd(Y1)^2+sd(Y2)^2+sd(Y3)^2+sd(Y4)^2+sd(Y5)^2)
=sqrt(3*1.1^2+5*1.43^2)
=3.72 <2 points>

HW_5_4 <10 points>
Omitted here.