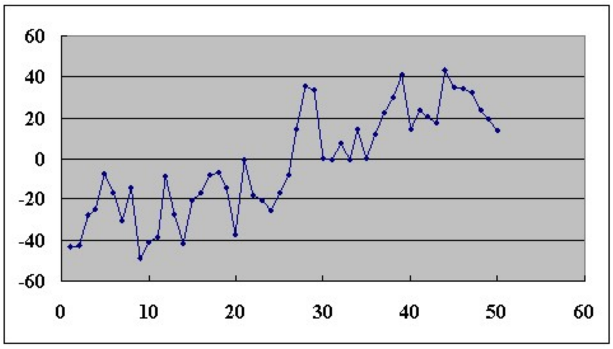
STATISTICS

Homework 5 (Due 5/10/2019)

1. The data in the following table is a subseries of an AR(2) process –
   $x(t) = 0.62x(t-1) + 0.25x(t-2) + \varepsilon(t) \quad \varepsilon(t) \sim iid\ N\left(0, \sigma_\varepsilon^2 = 225\right)$

t	1	2	3	4	5	6	7	8	9	10
0	-43.5615	-42.5594	-27.5556	-24.8670	-7.6353	-16.6590	-30.3863	-13.9924	-48.8766	-40.9939
10	-38.2022	-8.6007	-26.9491	-41.7347	-20.3306	-16.6453	-8.1240	-6.7680	-14.3989	-37.0406
20	-0.8470	-18.0262	-20.5075	-25.0793	-16.8221	-7.7765	14.4545	35.3056	33.6065	-0.1143
30	-0.5319	7.5514	-0.4998	14.1964	-0.1449	11.8305	22.5281	29.9923	40.7217	14.2982
40	23.3038	20.4037	17.2357	43.1575	34.5701	33.7827	31.9995	23.4265	18.9954	13.7399

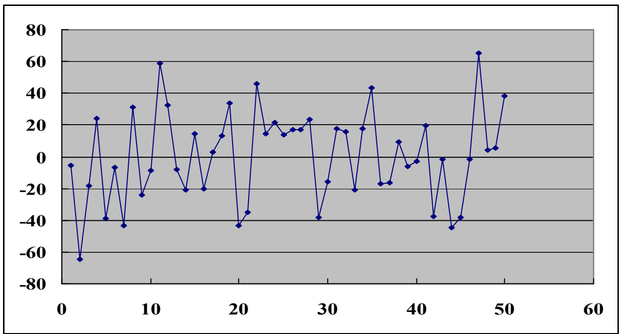


(1) Calculate the mean and standard deviation of $x(t),\ t = 1,2,\cdots,50.$

(2) The theoretical values of mean and standard deviation of X are respectively 0 and 27.5318. Compare these theoretical values against their estimates obtained in (1).

2. The following is a random sample from a normal distribution of mean 0 and standard deviation 27.5318.

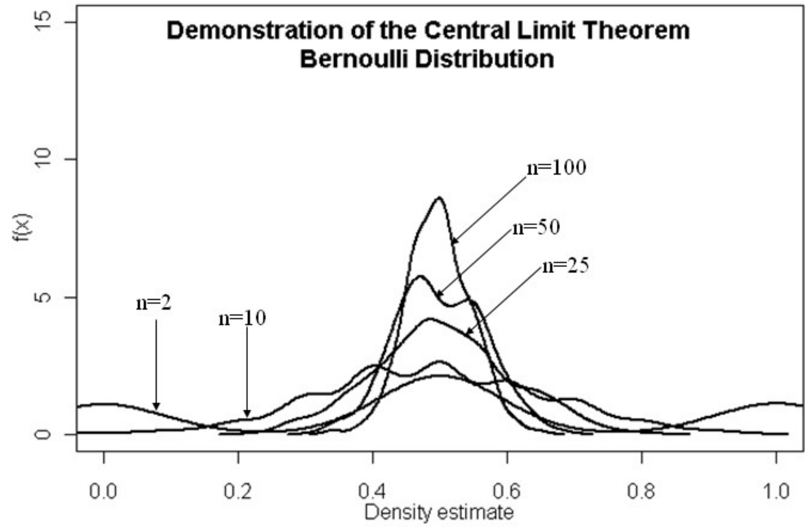
t	1	2	3	4	5	6	7	8	9	10
0	-5.407	-64.317	-18.434	24.298	-38.989	-6.800	-43.550	31.181	-23.789	-8.363
10	59.050	32.460	-8.182	-20.621	14.489	-20.165	2.946	13.432	33.898	-43.162
20	-35.052	46.106	14.658	21.340	13.584	16.901	16.977	23.161	-38.092	-16.026
30	17.982	15.979	-20.913	17.966	43.496	-17.282	-16.295	9.455	-6.092	-3.146
40	19.561	-37.615	-1.747	-44.600	-38.186	-1.586	65.115	4.176	5.509	38.330



(1) Calculate the sample mean and sample standard deviation.

(2) Are the sample estimates of mean and standard deviation in this problem closer to the theoretical values than the estimates in Problem 1?

3. Develop your own R code to demonstrate the central limit theorem for the sample mean of iid (independently, identically distributed) random variables with Bernoulli distribution (p=0.5). [Note: The R code should generate a figure similar to the following.]



4. Suppose that $\overline{X}_1$ and $\overline{X}_2$ are means of two random samples of size n from a population with variance σ^2. Determine n such that the probability will be about 0.01 that the two sample means will differ by more than σ. [Hint: Consider $Y = \overline{X}_1 - \overline{X}_2$.]

5. A research worker wishes to estimate the mean of a population using a sample large enough that the probability will be 0.95 that the sample mean will not differ from the population mean by more than 25% of the standard deviation. How large a sample should he take?

6. Let X_1 and X_2 be a random sample from $N(0,1)$.

(1) What is the distribution of $(X_2 - X_1)/\sqrt{2}$?

(2) What is the distribution of $(X_1 + X_2)^2/(X_2 - X_1)^2$?

(3) What is the distribution of $(X_1 + X_2)/\sqrt{(X_1 - X_2)^2}$?

7. Give an example of random sample that can be observed in your daily life and explain why you think they form a random sample.