Introduction to Probability, Statistics and Data Handling	**Confidence intervals**
Tutorial 7	

Discussion on confidence intervals:
Suppose you want to know what the **true mean value μ** of the population. You can take a sample and calculate the mean of the **sample $\overline{X}$ and the standard deviation s**. This the **point estimation** of the unknown μ and unknown σ of the population.
You can also calculate the **interval** which includes your unknown μ and unknown σ. It is called **confidence interval** and it with a given probability (**confidence level**) include μ and σ, unknown population parameters.
The confidence interval is created for an unknown population parameter like the population mean, μ.
Confidence intervals for some parameters have the form:

(point estimate – margin of error, point estimate + margin of error).

The margin of error depends on the confidence level or percentage of confidence and the standard error of the mean.
The methods of confidence level determination depend on the probability distribution it has and the size of the sample. Pay attention and practice!

1. *Interval estimation in large and small samples. Normal distribution.*
   a) A random sample of 120 students from a large university yields mean grade 2.7 with **population standard deviation σ =0.5.** Construct a 90% , 95% and 98% confidence interval for the mean grade of all students at the university. Assume that the population from which the sample is taken has a normal distribution of grades.
   b) Solve the above problem when the number of students is only 12
   c) How many students we need to randomly surveyed to be 95% confident that sample mean grade is within 2 of the true population grade?
   d) Solve b) but assume that you have **calculated the standard deviation s from the sample** and s =0.5. You need to use **t-Student distribution**.
2. *Interval estimation in large samples. Unknown distribution.*
   Suppose that an accounting firm does a study to determine the time needed to complete one person's tax forms. It randomly surveys 100 people. The sample mean is 23.6 hours. There is a known standard deviation of 7.0 hours.
   a) Construct a 90% confidence interval for the population mean time to complete the tax forms.
   b) If the firm wished to increase its level of confidence and keep the error bound the same by taking another survey, what changes should it make?
   c) If the firm did another survey, kept the error bound the same, and only surveyed 49 people, what would happen to the level of confidence? Why?
   d) Suppose that the firm decided that it needed to be at least 99% confident of the population mean length of time to within one hour. How would the number of people the firm surveys change? Why?
3. *Interval estimation for a population proportion:*
   a) Suppose that a market research firm is hired to estimate the percent of adults living in a large city who have cell phones. Five hundred randomly selected adult residents in this city are surveyed to determine whether they have cell phones. Of the 500 people surveyed, 421 responded yes - they own cell phones. Using a 95% confidence level, compute a confidence interval estimate for the true proportion of adult residents of this city who have cell phones.
   b) Suppose an internet marketing company wants to determine the current percentage of customers who click on ads on their smartphones. How many customers should the company survey in order to be 90% confident that the estimated proportion is within five percentage points of the true population proportion of customers who click on ads on their smartphones?