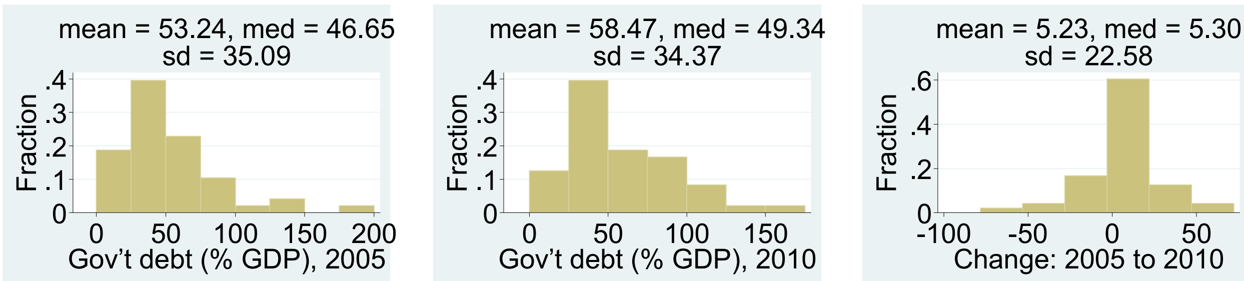
# ECO220Y: Homework 8

**Required Exercises:** Chapter 9: 7, 9, 11, 35, 37, 41, 43, 45, 47, 49, 61, 62 – 64

**Required Problems:**

**(1)** Below are three graphs – using data from the World Bank for 48 countries – from Slide 24 of Lecture 3. What is the coefficient of correlation between debt as a percent of GDP in 2005 and debt as a percent of GDP in 2010? Figure out the exact value and explain how you know the sign of the correlation must be what it is (i.e. either positive or negative).




**(2)** Give a numeric example where Y, which is defined as the sum of two random variables Y = W + X, has a smaller variance than either W or X.

**(3)** You are told that 12 percent of airline delays are caused by mechanical issues with the airplane. In a random sample of 25 flights, 5 (20%) are delayed by mechanical issues. What is the chance such a high number are delayed for a mechanical reason if the original claim is true? Do you think that your sample is statistically plausible? [Hint: The question is *not* asking for the chance exactly 5 out of 25 are delayed. "Such a high number" means that high or higher.]

**(4)** In a survey of adult GTA residents one question asks "Do you support Mayor John Tory? Yes or No." For all parts below suppose that in truth exactly 60% of the entire adult population of GTA residents support Tory.

> **(a)** Suppose random sample of 10 Toronto residents is asked the survey question. Draw a *fully labeled graph* of the distribution of the random variable recording the number of people who answer "Yes" to the survey.
>
> **(b)** What is the probability that the percent of the sample in Part (a) supporting Tory is at least 50 percent but not more than 70 percent?

**(5)** Which two parameters affect the probability of any particular number of successes in a Binomial experiment?

**(6)** For a Binomial random variable with $n = 14$ and $p = 0.5$, what is the probability of obtaining a value of $X$ within one standard deviation of the mean?

**(7)** A very large discount store is filled with items that have one of five different prices. 48.23 percent of the items in the store cost 99 cents, 38.58 percent cost \$1.99, 11.57 percent cost \$2.99, 1.54 percent cost \$3.99, and 0.08 percent cost \$4.99. Suppose you walked into the store and selected one item at random.

> **(a)** Define a random variable X and write down its probability distribution. Is it continuous or discrete?
>
> **(b)** Graph the probability distribution of X. Make sure to label the axes.
>
> **(c)** Find the mean. Interpret it in a way that a customer at this store could understand. Include the units of measurement.

**(d)** Find the s.d. and include the units of measurement.

**(e)** Create a variable $Z$ that is $X$ standardized. Write down its probability distribution. Interpret one of the negative values of $Z$. Interpret one of the positive values of $Z$. What are the units of measurement?

**(f)** Graph the probability distribution of $Z$. How does it compare with the graph for Part **(b)**?

**(8)** Suppose you have 1 year of data for a restaurant. It has 365 observations (records) corresponding to each day in that year. There are three variables recording the morning shift sales ($'s), the afternoon shift sales ($'s) and the evening shift sales ($'s). The mean and standard deviation of daily sales for each shift are:

Morning: $1100 and $200; Afternoon: $1300 and $150; Evening: $1600 and $400

Answer this multiple-choice question and include an explanation of the correct and incorrect replies. For total daily sales (three shifts combined), what is the mean and standard deviation?

**(A)** mean = $4000 and s.d. = $750
**(B)** mean = $4000 and s.d. = $471.70
**(C)** mean cannot be determined but s.d. = $750
**(D)** mean = $4000 but s.d. cannot be determined

**(9)** Consider again the OECD data discussed in Lecture 8.

```
correlate fem_emp_2006 fem_emp_2012 male_emp_2006 male_emp_2012, format;

(obs=34)
             | fem~2006 fem~2012 mal~2006 mal~2012
-------------+------------------------------------
fem_emp_2006 |   1.0000
fem_emp_2012 |   0.9482   1.0000
male_em~2006 |   0.5904   0.4948   1.0000
male_em~2012 |   0.4758   0.6106   0.6498   1.0000

summarize fem_emp_2006 fem_emp_2012 male_emp_2006 male_emp_2012, format;

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
fem_emp_2006 |        34       58.92       11.45      22.75      80.88
fem_emp_2012 |        34       60.05       10.43      28.73      77.88
male_em~2006 |        34       74.43        6.31      60.88      88.05
male_em~2012 |        34       72.23        6.34      60.20      85.18
```

**(a)** What is the mean and the s.d. of the *change* in the male employment rate from 2006 to 2012 across OECD countries? Include the units of measurement of each.

**(b)** Select true or false for each.

**i)** For $D = [M_{12} - M_{06}]$, the mean of D is positive
**ii)** For $D = [M_{12} - M_{06}]$, the variance of D is larger than $V[M_{12}] + V[M_{06}]$
**iii)** For $D = [M_{12} - M_{06}]$, the variance of D is affected by the sign of the correlation
**iv)** For $D = [M_{12} - M_{06}]$, the s.d. of D is $SD[M_{12}] + SD[M_{06}] - 2rSD[M_{12}]SD[M_{06}]$
**v)** For $D = [M_{12} - M_{06}]$, the variance of D is the same as the variance of $[M_{12} + M_{06}]$
**vi)** For $D = [M_{12} - M_{06}]$, the s.d. of D is not equal to $SD[M_{12}] - SD[M_{06}]$

**(10)** Consider doing a Binomial Experiment and defining a new random variable $Y$ as the fraction of successes in the $n$ trials. For example, if there are 3 successes in 5 trials, $Y$ would be 0.60 whereas the Binomial random variable $X$ would be 3. Is $Y$ a discrete random variable? Explain. What are the mean and variance of $Y$?

**(11)** Draw a graph of a Binomial distribution if n = 6 and p = 0.2. Carefully label it. What is the mean and variance? Explain how the shape would change if n = 600? What would be the new mean and variance?

**(12)** Suppose the probability distribution of the number of sales a sales person makes in a day is:

| x | p(x) |
|---|---|
| 0 | 0.4 |
| 1 | 0.5 |
| 2 | 0.1 |

**(a)** What is the expected total sales for a week (with 5 business days)?

**(b)** If the number of sales made each day is independent of all other days, what is the s.d. of total sales for a week (with 5 business days)?