Mid-Term Review

Mid-Term

- Saturday June 17, 9.00a.m CT
- Submit on Tuesday June 20, 11.00p.m. CT
- Text book, Live session notes, Labs, HWs, asynchronous videos (week1-week5).
- Office hour : Friday 6.30p.m.-7.30p.m. CT

Mid-Term

- Multiple choice questions
- Sampling Distribution calculations
- SAS/R
- Excel
- Conceptual Questions

Lab 04

```
data strsizes;
input stratum _total_;
datalines;
1 N1
2 N2
3 N3
4 N4
5 N5
;
proc surveymeans data = neysample sum clsum total = strsizes
mean sum CLSUM;
var sales;
weight SamplingWeight;
strata stratum;
title "Neyman allocation";
run;
```

Yes!

Read CSV file

```
data industry2;
infile "C:\MSDS 6370 FALL2016\Live session
04\MSDS6370Lab4.csv"
dlm="," firstobs=2;
input  industry UnitId mos stratum Sales  ;
run;
```

HW4

1) Sampling distribution of a finite population

Name	GPA
A	3.5
B	2.7
C	3.1
D	3.2
E	2.4

How many possible samples of size 3?

$$\binom{5}{3} = \frac{5!}{(2)!3!} = \frac{20}{2} = 10$$

HW4-10 possible samples

Sample id	sample members	Sample mean
1	ABC	3.10
2	ABD	3.13
3	ABE	2.87
4	ACD	3.27
5	ACE	3.00
6	ADE	3.03
7	BCD	3.00
8	BCE	2.73
9	BDE	2.77
10	CDE	2.90

HW4-Sampling Distribution of sample means

Sample mean	Probability
2.73	0.1
2.77	0.1
2.87	0.1
2.90	0.1
3.00	0.2
3.03	0.1
3.10	0.1
3.13	0.1
3.27	0.1

What is the mean of the sampling distribution of sample means?

2.98

MEAN OF A DISCRETE RANDOM VARIABLE

Suppose that X is a discrete random variable whose distribution is

Value of X	x_1	x_2	x_3	$\cdots$	x_k
Probability	p_1	p_2	p_3	$\cdots$	p_k

To find the **mean** of X, multiply each possible value by its probability, then add all the products:

$$\mu_X = x_1p_1 + x_2p_2 + \cdots + x_kp_k$$
$$= \sum x_ip_i$$

VARIANCE OF A DISCRETE RANDOM VARIABLE

Suppose that X is a discrete random variable whose distribution is

Value of X	x_1	x_2	x_3	$\cdots$	x_k
Probability	p_1	p_2	p_3	$\cdots$	p_k

and that μ is the mean of X. The **variance** of X

$$\begin{aligned}\sigma_X^2 &= (x_1 - \mu_X)^2 p_1 + (x_2 - \mu_X)^2 p_2 + \cdots + (x_k - \mu_X)^2 p_k \\ &= \sum (x_i - \mu_X)^2 p_i\end{aligned}$$

The **standard deviation** σ_X of X is the square root of the variance.

HW4

What is the standard error of the sampling distribution of sample means?

0.158

Sample mean	Probability	
2.73	0.1	$(2.73 - 2.98)^2\ 0.1$
2.77	0.1	$(2.77 - 2.98)^2\ 0.1$
2.87	0.1	$(2.87 - 2.98)^2\ 0.1$
2.90	0.1	$(2.90 - 2.98)^2\ 0.1$
3.00	0.2	$(3.00 - 2.98)^2\ 0.2$
3.03	0.1	$(3.03 - 2.98)^2\ 0.1$
3.10	0.1	$(3.10 - 2.98)^2\ 0.1$
3.13	0.1	$(3.13 - 2.98)^2\ 0.1$
3.27	0.1	$(3.27 - 2.98)^2\ 0.1$

$$\sqrt{Total}$$

Alternative Way

$$\sigma_{\bar{y}} = \sqrt{\frac{S^2}{n}\left(1 - \frac{n}{N}\right)},$$

where $S^2 = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \bar{Y})^2$

Name	GPA
A	3.5
B	2.7
C	3.1
D	3.2
E	2.4

$$S^2 = 0.187\ (Var.s)$$

$$\sigma_{\bar{y}} = \sqrt{\frac{0.187}{3}\left(1 - \frac{3}{5}\right)} = 0.158$$