Assignment #2, Due 7/24(25)/2024

1. In class, we use the court example to explain the null and alternative hypotheses. Based on the case of People vs. Collins (1969), discuss the following topics:

   (a) What are the null hypothesis and the alternative hypothesis?

   →虛無假設H_0：被告無罪 vs. H_a：被告有罪

   (b) Explain your reasons for the choice of the null hypothesis.

   →假設檢定的結果只有拒絕或不拒絕H_0，如果H_0是被告有罪，必須找證據駁斥，但除非找到真兇否則無法證明被告無罪。

   (c) Do you think the testimony of eyewitnesses can be used to charge the couples?

   →首先，目擊證人的證詞有資料品質的問題，再加上機率相乘必須滿足獨立事件，但明顯並不成立，還有事件發生機率不等同於p-value，不能將其解讀為被告無罪的可能性。

2. **Median Annual Income in Chicago.** The median annual income for families living in the United States is $56,200. Annual incomes in thousands of dollars for a sample of 50 families living in Chicago, Illinois, are shown. Use the sample data to see if it can be concluded that the families living in Chicago have a median annual income greater than $56,200. Use α = .05. What is your conclusion?

66.3	60.2	49.9	75.4	73.7
65.7	61.1	123.8	57.3	48.5
74	146.3	92.2	43.7	86.9
59.7	64.2	56.2	48.9	109.6
39.8	60.9	79.7	42.3	52.6
60.9	43.5	61.7	54.7	95.2
70.4	43.8	57.8	83.5	56.5
51.3	42.9	87.5	43.6	67.2
48.7	79.1	61.9	53.4	56.2
57	49.6	109.5	42.1	74.6

   →虛無假設 H_0：中位數收入＝$56,200（或中位數收入≤＄56,200）vs. H_a：中位數收入>$56,200。48 位受訪者（扣除兩位數值為$56,200）中有 31 位收入大於$56,200，以常態分配近似二項分配，或是 N(24,12)近似 B(48, 0.5)，得出 $P(X \geq 31)＝P(Z > 2.02)＝0.0217$，無論是單尾或雙尾檢定都是拒絕 H_0。

3. A clothing manufacturer purchased some newly designed sewing machines in the hopes that production would be increased. The production records of a random sample of workers are shown below.

Worker	Old Machine	New Machine
1	28	36
2	36	40
3	27	25
4	25	32
5	38	30
6	36	32
7	40	40
8	29	28
9	32	35
10	28	33
11	20	26
12	32	31
13	32	23
14	32	34
15	36	36

Use the Wilcoxon signed-rank test to determine whether the new machines have significantly increased production. Use a .05 level of significance.

→由題意設定本題為單尾檢定，其中 H_0 為新舊機器產量無差異。先計算兩種機器的 Wilcoxon signed-rank 的檢定數值，將焦點放在 T^-，亦即新機器產量比較多（扣除產量相同者）：

Worker	Old Machine	New Machine	Difference	Rank	T^-
1	28	36	-8	-11.5	11.5
2	36	40	-4	-6.5	6.5
3	27	25	2	3.5	
4	25	32	-7	-10	10
5	38	30	8	11.5	
6	36	32	4	6.5	
8	29	28	1	1.5	
9	32	35	-3	-5	5
10	28	33	-5	-8	8
11	20	26	-6	-9	9
12	32	31	1	1.5	
13	32	23	9	13	
14	32	34	-2	-3.5	3.5

計算可得 $T^-=53.5$，而理論期望值為 $91/2=45.5$、變異數＝204.75，因此檢定量為 $(53.5-45.5)/\sqrt{204.75}=0.5591$，計算可得 p-value＝0.2880，不拒絕 H_0。

4. Quality of Teaching Assessments. A student organization surveyed both current

students and recent graduates to obtain information on the quality of teaching at a particular university. An analysis of the responses provided the following teaching-ability rankings. Do the rankings given by the current students agree with the rankings given by the recent graduates? Use $\alpha = .10$ and test for a significant rank correlation.

Professor	Current Students	Recent Graduates
1	4	6
2	6	8
3	8	5
4	3	1
5	1	2
6	2	3
7	5	7
8	10	9
9	7	4
10	9	10

→本題以Spearman相關係數探討在學學生、畢業生給予老師的排名是否有關，虛無假設H_0是兩者無關。計算得出相關係數值為0.7697，標準誤的估計值為 $\sqrt{\frac{1}{9}} = \frac{1}{3}$，因此檢定值為2.3091，或是雙尾檢定的p-value＝0.0209。因此拒絕H_0，也就是認定相關係數顯著異於0，亦即在學學生、畢業生給予老師的排名順序，兩者之間有關聯。

5. **Production Process Temperature.** Temperature is used to measure the output of a production process. When the process is in control, the mean of the process is $\mu = 128.5$ and the standard deviation is $\sigma = .4$.

(a) Construct the $\overline{x}$ chart for this process if samples of size 6 are to be used.

→假設期望值及變異數已知，因此$\overline{x}$ chart 的管制界線為$(\mu-3\sigma_{\overline{x}}, \mu+3\sigma_{\overline{x}})＝$ (128.01, 128.99)，其中$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{6}} = 0.1633$。

(b) Is the process in control for a sample providing the following data?

128.8　　128.2　　129.1　　128.7　　128.4　　129.2

→六筆資料平均數為 128.8，落在管制界線之內，製程在控制之中。

(c) Is the process in control for a sample providing the following data?

129.3　　128.7　　128.6　　129.2　　129.5　　129.0

→六筆資料平均數為 129.3、超出管制範圍，因此製程失控。

6. **Airline Call Center.** An airline operates a call center to handle customer questions and complaints. The airline monitors a sample of calls to help ensure that the service being provided is of high quality. Ten random samples of 100 calls each were monitored under normal conditions. The center can be thought of as being in control when these 10 samples were taken. The number of calls in each sample not resulting in a satisfactory resolution for the customer is as follows:

$$4 \quad 5 \quad 3 \quad 2 \quad 3 \quad 3 \quad 4 \quad 6 \quad 4 \quad 7$$

(a) What is an estimate of the proportion of calls not resulting in a satisfactory outcome for the customer when the center is in control?

→因為樣本數夠大，可用常態分配近似，在十次100個觀察值的抽樣中，估計有$\hat{p}$＝41/1000＝0.041的機會發生失誤。(註：雖然$np < 5$，但數值接近5，仍然可代入常態分配的假設。

(b) Construct the upper and lower limits for a *p* chart for the manufacturing process, assuming each sample has 100 calls.

→*p* chart的管制範圍為$(\hat{p} - 3\sigma_{\hat{p}}, \ \hat{p} + 3\sigma_{\hat{p}})$＝(0.041–3×.0198, 0.041+3×.0198)＝(0, 0.1005)，需調整下限至0。

(c) With the results of part (b), what conclusion should be made if a sample of 100 has 12 calls not resulting in a satisfactory resolution for the customer?

→100個觀察值中有12次瑕疵，判定為失控。

(d) Compute the upper and lower limits for the *np* chart.

→*np* chart只是將瑕疵機率轉換成瑕疵次數，管制範圍為(0, 10.05)。