ISYS3401
Information Technology Evaluation

Week 4 Lecture

Dr Vincent Pang
Vincent.Pang@sydney.edu.au

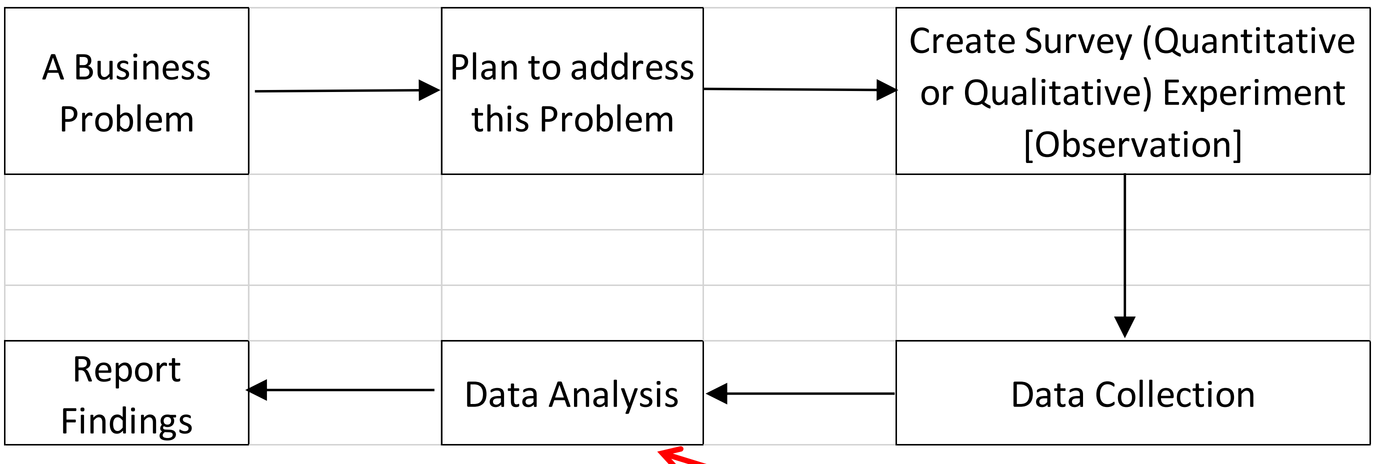
Agenda

- Recap
- $\chi 2$ test
- McNemar's Test
- t-test

Recap …

In Week 1, you were asked by Mr Apple to conduct UX Experience (for $10 million).

Next week, we will start doing survey questions (see Chapter 6)



Last week, we did chi-square test and t-test, will give you the answers this week… Plus McNemar’s test

Significance Testing
(Hypothesis Testing)

1. Formulate a null hypothesis (H_0) that there is no difference (or no effect)
2. Collect data to test the hypothesis
3. Calculate the probability (P) of these or more extreme data occurring if the null hypothesis is true
4. If **P is large**, the null hypothesis cannot be rejected. This does not necessarily mean that the null hypothesis is true. The result is said **to be not statistically significant.**
5. If **P is small**, we **reject the null hypothesis**, and conclude that there is an effect which is said **to be statistically significant**.

For T-Test

- H_0: There is no difference between Sample A and Sample B
  - If t-value is lower than the Critical Value, then do not Reject
  - If t-value is higher than the Critical Value, then Reject

Testing Methods (σ unknown)

For Evaluation Studies

	One Sample	
	Continuous variable	**Binomial variable**
One sample	Estimate ***s*** from the sample and use the student's ***t***	Normal approximation to Binomial (equivalent to χ^2 test)

	Two Samples (to be compared)	
	Continuous variable	**Binomial variable**
Paired samples	Normally distributed (approximately) Paired ***t*** test	McNemar's test
Independent samples	Normally distributed (approximately) 2-sample ***t*** test	2 samples χ^2 test (2 x 2 table) Comparison of 2 proportions

Chi-Square Test

- The chi-square (independence) test is a procedure for testing if two categorical variables are related in some population. [SPSS Statistics]

Chi-Square Formula

Two Formulae you can use:

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

	O	E	O-E	|O-E|-½
Prefer N	10	6	4	3.5
Prefer E	2	6	-4	3.5
	12	12		

$$\chi^2 = \frac{(ad-bc)^2 \times n}{n_1 n_2 m_1 m_2}$$

	Binomial category (e.g. outcome)		
	A	B	Total
Group 1	a	b	n_1
Group 2	c	d	n_2
Total	m_1	m_2	n

Example using First Formula (1)

Example: Cross-over trial to compare a new web design (N) an the existing website (E) for user acceptance (on ease of use)

- N=12 with
  - 10 students prefers N over E
  - 2 students prefers E over N
- Null hypothesis (H_0): N and E are equally accepted

Example using First Formula (2)

chi-square (χ^2) statistic

- Using the First Formula to calculate:

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

where:

- ***O*** represents the observed count
- ***E*** represents the expected count under the assumption that the null hypothesis is true

- the Expected numbers under the null hypothesis are that **6 would prefer N** and **6 would prefer E**.
- It is preferable to include a continuity correction by moving each observed value ½ closer towards its expected value.
- Applying the formula and outline the numbers in the table.

	O	E	(O-E)	(O-E-(1/2))	$(O-E)^2$	$(O-E)^2/E$
Prefer N	10	6	4	3.50	12.25	2.04
Prefer E	2	6	-4	3.50	12.25	2.04
	12	12				4.08

$$\chi^2 = \frac{3.5^2}{6} + \frac{3.5^2}{6} = 4.08$$

Example using First Formula (3)

chi-square (χ^2) statistic

- χ^2 is approximately distributed as a chi-squared variable with 1 degree of freedom
- The critical value of χ^2 (1df) distribution for P=0.05 is 3.84.
- Since 4.08 > 3.84, P < 0.05 for this example
- Hence we conclude that N is more acceptable than E (P < 0.05)

TABLE C: χ^2 CRITICAL VALUES

	Tail probability p										
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001
1	1.32	1.64	2.07	2.71	$\boxed{3.84}$	5.02	5.41	6.63	7.88	9.14	10.83
2	2.77	3.22	3.79	4.61	5.99	7.38	7.82	9.21	10.60	11.98	13.82
3	4.11	4.64	5.32	6.25	7.81	9.35	9.84	11.34	12.84	14.32	16.27
4	5.39	5.99	6.74	7.78	9.49	11.14	11.67	13.28	14.86	16.42	18.47
5	6.63	7.29	8.12	9.24	11.07	12.83	13.39	15.09	16.75	18.39	20.51
6	7.84	8.56	9.45	10.64	12.59	14.45	15.03	16.81	18.55	20.25	22.46

Example using Second Formula (1)

- The situation arises when the proportion of individuals with a binary outcome is obtained in two groups. The two groups are independent, i.e. there is no pairing of members in one group with members in the other.
- The two independent groups could be defined in different ways, examples:
  - Random allocation of each individual in the study to either new solution or control (e.g. drug treatment versus placebo)
  - One group comprises males, the other females
- Since each individual belongs to one of two groups, and to one of two binomial categories, this gives rise to four types of individual. The number of each type can be represented in a 2×2 table

	Binomial category (e.g. outcome)		
	A	B	Total
Group 1	a	b	n_1
Group 2	c	d	n_2
Total	m_1	m_2	n

Example using Second Formula (2)

- Suppose that in the population that has been sampled the probability that an individual is of type A is:
  - θ_1 for group 1
  - θ_2 for group 2
- Null hypothesis: Both samples are from a population in which the probability that an individual is of type A is θ, that is $\theta_1 = \theta_2$
- In a survey of 244 USYD graduates, 75 out of 124 male graduates never studied IT units and 62 out of 120 female graduates never studied IT units in their undergraduate studies. Is the proportion of "IT literacy" different for the two sexes?

Observed	**Binomial category (e.g. outcome)**		
	Never studied IT	Studied IT	Total
Male students	75 (a)	49 (b)	124
Female students	62 (c)	58 (d)	120
Total	137	107	244 (n)

Computing the Expected counts Using First Formula

Expected	**Binomial category (e.g. outcome)**		
	Never studied IT	Studied IT	Total
Male students	124 x 137 / 244 = 69.62 (a)	124 x 107 /244 = 54.38 (b)	124
Female students	120 x 137 / 244 = 67.38 (c)	120 x 107 / 244 = 52.62 (d)	120
Total	137	107	244 (n)

	Observed O	**Expected E**	**O – E**	**$(O-E)^2/E$**
a	75	69.62	5.38	0.416
b	49	54.38	-5.38	0.532
c	62	67.38	-5.38	0.430
d	58	52.62	5.38	0.550
Total	244	244.00	0	1.928

- From the survey of 244 USYD students it was found that 60% of male students and 52% of female students never studied IT. These percentages did not differ significantly (χ^2 = 1.93, 1df, P=0.16).

Example using Second Formula (3)

Observed	Binomial category (e.g. outcome)		
	Never studied IT	Studied IT	Total
Male students	75 (a)	49 (b)	124 (n_1)
Female students	62 (c)	58 (d)	120 (n_2)
Total	137 (m_1)	107 (m_2)	244 (n)

- For a 2×2 table, the χ2 test statistic can be calculated more easily using the formula

$$\chi^2 = \frac{(ad - bc)^2 \times n}{n_1 n_2 m_1 m_2}$$

$$= \frac{(75 \times 58 - 62 \times 49)^2 \times 244}{124 \times 120 \times 137 \times 107} = 1.93 \quad with \quad 1df$$

Example using First Formula (4)

chi-square (χ^2) statistic

- χ^2 is approximately distributed as a chi-squared variable with 1 degree of freedom
- The critical value of χ^2 (1df) distribution for P=0.05 is 3.84.
- Since 1.93 < 3.84, P > 0.05 for this example
- Hence, we conclude that there is no “IT literacy” different for the two sexes

TABLE C: χ^2 CRITICAL VALUES

	Tail probability p										
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001
1	1.32	1.64	2.07	2.71	$\boxed{3.84}$	5.02	5.41	6.63	7.88	9.14	10.83
2	2.77	3.22	3.79	4.61	5.99	7.38	7.82	9.21	10.60	11.98	13.82
3	4.11	4.64	5.32	6.25	7.81	9.35	9.84	11.34	12.84	14.32	16.27
4	5.39	5.99	6.74	7.78	9.49	11.14	11.67	13.28	14.86	16.42	18.47
5	6.63	7.29	8.12	9.24	11.07	12.83	13.39	15.09	16.75	18.39	20.51
6	7.84	8.56	9.45	10.64	12.59	14.45	15.03	16.81	18.55	20.25	22.46

Tutorial Question 3

In a study of the association between amount of internet surfing and being overweight, 290 staff members of USYD students were surveyed. 243 said they surfed at least one hour of internet each day, while the rest claimed to surf less than one hour per day. Of the 243, 106 were found to be overweight (body mass index > 25kg/m2); 14 of those surfing less than one hour per day were overweight. [Note: you need to define Null hypothesis first.]

- Analyse these data to examine the association between amount of internet surfing and being overweight.
- What is your conclusion?

TQ3: Calculate Chi-Square

A. H_0: There is no association between being overweight and surfing more than one hour of internet per day.

		Hours of internet surfing/day		
		<1	≥ 1	Total
Overweight	Yes	14	106	120
	No	33	137	170
	Total	47	243	290

Chi-squared test: $$\chi^2 = \frac{(14 \times 137 - 106 \times 33)^2 \times 290}{120 \times 170 \times 47 \times 243} = 3.11 \quad \textit{with } 1df$$ (P > 0.05)

TQ3: Conclusion

chi-square (χ^2) statistic

- χ^2 is approximately distributed as a chi-squared variable with 1 degree of freedom
- The critical value of χ^2 (1df) distribution for P=0.05 is 3.84.
- Since 3.11 < 3.84, null hypothesis cannot be rejected.

[Further test might be required.]

TABLE C: χ^2 CRITICAL VALUES

	Tail probability p										
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001
1	1.32	1.64	2.07	2.71	$\boxed{3.84}$	5.02	5.41	6.63	7.88	9.14	10.83
2	2.77	3.22	3.79	4.61	5.99	7.38	7.82	9.21	10.60	11.98	13.82
3	4.11	4.64	5.32	6.25	7.81	9.35	9.84	11.34	12.84	14.32	16.27
4	5.39	5.99	6.74	7.78	9.49	11.14	11.67	13.28	14.86	16.42	18.47
5	6.63	7.29	8.12	9.24	11.07	12.83	13.39	15.09	16.75	18.39	20.51
6	7.84	8.56	9.45	10.64	12.59	14.45	15.03	16.81	18.55	20.25	22.46

Tutorial Question 4

A case-control study was carried out to examine the association between open source software (OSS) adoption and formal position of CIO. There were 148 firms identified as cases (adopted OSS), and 537 firms were randomly selected as controls. 79 of the cases do not have a formal position as CIO and 69 do have a formal position as CIO in their organisation. For controls, 260 do not have a formal position as CIO and 277 do have a formal position as CIO in their organisation. [Note: you need to define Null hypothesis first.]

- Analyse these data to test if there is any evidence of an association between OSS adoption and formal position of CIO.
- What is your conclusion?

TQ4: Calculate Chi-Square

A. H_0: There is no association between OSS adoption and having a formal position of CIO.

	Formal position of CIO		
	Yes	No	Total
Cases	69	79	148
Control	277	260	537
Total	346	339	685

Chi-squared test: $$\chi^2 = \frac{(69\times 260 - 79\times 277)^2 \times 685}{148\times 537\times 346\times 339} = 1.14 \quad with\ 1df$$ (P > 0.05)

TQ4: Conclusion

chi-square (χ^2) statistic

- χ^2 is approximately distributed as a chi-squared variable with 1 degree of freedom
- The critical value of χ^2 (1df) distribution for P=0.05 is 3.84.
- Since 1.14 < 3.84, null hypothesis cannot be rejected.
- Thus, there is no evidence association between OSS adoption and having formal position of CIO in the organisation.

TABLE C: χ^2 CRITICAL VALUES

	Tail probability p										
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001
1	1.32	1.64	2.07	2.71	$\boxed{3.84}$	5.02	5.41	6.63	7.88	9.14	10.83
2	2.77	3.22	3.79	4.61	5.99	7.38	7.82	9.21	10.60	11.98	13.82
3	4.11	4.64	5.32	6.25	7.81	9.35	9.84	11.34	12.84	14.32	16.27
4	5.39	5.99	6.74	7.78	9.49	11.14	11.67	13.28	14.86	16.42	18.47
5	6.63	7.29	8.12	9.24	11.07	12.83	13.39	15.09	16.75	18.39	20.51
6	7.84	8.56	9.45	10.64	12.59	14.45	15.03	16.81	18.55	20.25	22.46

Testing Methods (σ unknown)

For Evaluation Studies

	One Sample	
	Continuous variable	**Binomial variable**
One sample	Estimate ***s*** from the sample and use the student's ***t***	Normal approximation to Binomial (equivalent to χ^2 test)

	Two Samples (to be compared)	
	Continuous variable	**Binomial variable**
Paired samples	Normally distributed (approximately) Paired ***t*** test	McNemar's test
Independent samples	Normally distributed (approximately) 2-sample ***t*** test	2 samples χ^2 test (2 x 2 table) Comparison of 2 proportions

McNemar's test

- McNemar's test is equivalent to the χ^2 test.
- However, it is used for pairs matched. For instance, you are looking at age and gender, i.e. the pair variables, age and gender, belong to the same person, i.e. the variables/categories are not independent from each other.
- If *r is the number in one category and s is the number* in the other category, then McNemar's formula to compute is:

$$\chi^2 = \frac{(|r-s|-1)^2}{(r+s)}$$

- And is distributed as chi-squared with 1df
- Example (given r=10, s=2): $\chi^2 = \frac{(|10-2|-1)^2}{(10+2)} = \frac{7^2}{12} = 4.08$

McNemar's Test (Example)

- Null hypothesis: The proportion of IT graduates who agree that learning programming skill is important, is the same as the proportion who agree that analytical skill is important.

		Analytical skills important		
		Yes	No	Total
Programming skills important	Yes	144	2	146
	No	40	8	48
	Total	184	10	194

- Analytical skill is important: $p_1 = (144+40)/194 = 0.948$
- Programming skill is important: $p_2 = (144+2)/194 = 0.753$
- Taking the difference between the two proportions, the 144 graduates who agree that both are important have cancelled out. The statistical significance of the difference depends only on the number of discordant pairs:

$$p_1 - p_2 = \frac{144+40}{194} - \frac{144+2}{194} = \frac{40-2}{194} = 0.196$$

McNemar's Test (Example)

Applying the formula (r=40 and s=2):

$$\chi^2 = \frac{(|r-s|-1)^2}{(r+s)}$$

$$= \frac{(|40-2| - 1)^2}{(40+2)}$$

$$= \frac{1369}{42} = 32.59$$

Conclusion: There is very strong evidence that more IT graduates believe that analytical skill is more important than programming skill (χ^2 = 32.6, 1*df*)

T-Test

See Separate Document

Next Week

Read:

Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics, by William Albert, Thomas Tullis, ***Chapter 6, Self-Reported Metrics***