Q1 (a) Proportion of vaccinated adults is claimed to be 0.53. For testing this, what is usually taken as null hypothesis?

$X_1, X_2, \ldots, X_n \sim \text{Bernoulli}(p)$

$H_0: \; p = 0.53$

Q1 (b) If you suspect the fraction to be too high, what is your alternative hypothesis?

$H_A: \; p < 0.53$

rejection region should not include "null"

Q1 (c) In a survey, the following is found. What will your test be?

$n = 31$

$T = X_1 + \cdots + X_n = 21$

$\sim \text{Binomial}(n, p)$

Test: Reject H_0 if $T < c$

$c = 0.53 \times 31 - 1.645 \times 0.4991 \sqrt{31} = 11.85$

@ significance level $= 0.05$, $n = 31$

Q1 (d) Suppose you desire a type-I error probability of 0.05 and power of 0.95 against the alternative that the fraction if 0.45. How many samples do you need? Use the normal approximation to the Gaussian distribution.

$$\alpha = P(T < c \mid p = 0.53) \approx F_Z\left(\frac{c - 0.53n}{\sqrt{n \times 0.53 \times 0.47}}\right) = 0.05 \Rightarrow c = 0.53n - 1.645 \times 0.4991\sqrt{n}$$

$$\text{Power} = P(T < c \mid p = 0.45) \approx F_Z\left(\frac{c - 0.45n}{\sqrt{n \times 0.45 \times 0.55}}\right) = 0.95 \Rightarrow c = 0.45n + 1.645 \times 0.4975\sqrt{n}$$

Equating, $0.08n = 1.6394\sqrt{n} \Rightarrow n = 419.94$

Q2 (a) Average age of students is claimed to be 24. For testing this, what is usually taken as null hypothesis? $X_1, \ldots, X_n \sim N(\mu, \sigma^2)$

$H_0: \mu = 24$

Q2 (b) If you suspect the above to be incorrect, what is your alternative hypothesis?

$H_A: \mu \neq 24$

Q2 (c) In a survey, the following is found. What will your test be? σ^2: unknown

$n = 33$
$T = \bar{X} = 31$
$S^2 = 14^2$

Two-sided, t-test
Reject H_0 if $|T - 24| > c$

Q2 (d) At a significance level of 0.05, what is your conclusion? $\frac{T-24}{S/\sqrt{n}} \sim t_{n-1}$

$P(|T-24| > c \mid \mu = 24) \approx 2F_{t_{n-1}}\left(\frac{-c}{S/\sqrt{n}}\right)$

Put c = test statistic from survey & reject H_0 if value is < 0.05

Q3 (a) Average Quiz 2 marks in Stats II and Math II are equal. For testing this, what is usually taken as null hypothesis? Assume standard deviation to be 23 for both subjects.

$X_1, \dots, X_{n_1} \sim N(\mu_1, \sigma^2)$, $Y_1, \dots, Y_{n_2} \sim N(\mu_2, \sigma^2)$ $\sigma^2 = 23^2$

Stats $\qquad H_0 : \mu_1 = \mu_2 \qquad$ math

Q3 (b) If you suspect that Stats II has higher average, what is your alternative hypothesis?

$H_A : \mu_1 > \mu_2$

Q3 (c) In a survey, the following is found. What will your test be?

$n_1 = 20 \qquad n_2 = 20 \qquad$ Two sample Z-test

$\bar{X} = 76.8 \qquad \bar{Y} = 79.95 \qquad$ Reject H_0 if $\bar{X} - \bar{Y} > c$

Q3 (d) At a significance level of 0.05, what is your conclusion?

$$P(\bar{X} - \bar{Y} > c \mid \mu_1 = \mu_2) = 1 - F_Z\left(\frac{c}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}}\right) = 1 - F_Z\left(\frac{c/23}{\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}\right)$$

Put $c = \bar{X} - \bar{Y}$ from survey. If value is below 0.05, reject H_0

Q4 (a) Standard deviation of Quiz 2 marks in Stats II is 23. For testing this, what is usually taken as null hypothesis? $X_1, \ldots, X_n \sim N(\mu, \sigma^2)$

$H_0 : \sigma = 23$

Q4 (b) If you suspect that it is actually higher, what is your alternative hypothesis?

$H_A : \sigma > 23$

Q4 (c) In a survey, the following is found. What will your test be?

$n = 20$

$s^2 = 14^2$

χ^2 test for variance

Reject H_0 if $S > c$

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$$

Q4 (d) At a significance level of 0.05, what is your conclusion?

$$P(S > c \mid \sigma = 23) = 1 - F_{\chi^2_{n-1}}\left(\frac{(n-1)c^2}{23^2}\right)$$

For $c = S$ from survey, if above value is below 0.05, reject H_0

Q5 (a) Standard deviation of Quiz 2 marks in Stats II and Math II are equal. For testing this, what is usually taken as null hypothesis?

$X_1, \dots, X_{n_1} \sim N(\mu_1, \sigma_1^2)$ $\quad$ $Y_1, \dots, Y_{n_2} \sim N(\mu_2, \sigma_2^2)$

Stats $\quad$ $H_0 : \sigma_1^2 = \sigma_2^2$ $\quad$ Math

Q5 (b) If you suspect that Stats II's variance is higher, what is your alternative hypothesis?

$H_A : \sigma_1^2 > \sigma_2^2$

Q5 (c) In a survey, the following is found. What will your test be?

$n_1 = 20$ $\quad$ $n_2 = 20$

$S_X^2 = 19$ $\quad$ $S_Y^2 = 14$

Two sample F-test

Reject H_0 if $T = \frac{S_X^2}{S_Y^2} > 1 + c_R$

Q5 (d) At a significance level of 0.05, what is your conclusion?

$$P\left(\frac{S_X^2}{S_Y^2} > 1 + c_R \,\middle|\, \sigma_1 = \sigma_2\right) = 1 - F_{F(n_1-1, n_2-1)}(1 + c_R)$$

Put $1 + c_R = S_X^2 / S_Y^2$ from survey. If value < 0.05, reject H_0.

# medals	0	1	2	3	4	5	6	>6
fit	(1-p)/2	(1-p)/2	p/6	p/6	p/6	p/6	p/6	p/6
freq	4	5	1	1	0	0	1	0
expected	4.5	4.5	0.5	0.5	0.5	0.5	0.5	0.5

$k=8$

Q6 (a) Estimate p using ML method.

$$L \propto (1-p)^9 p^3 \Rightarrow \hat{p}_{ML} = \frac{3}{9+3} = \frac{1}{4}$$

Q6 (b) Is this a good fit using the chi-square test? What is the P-value?

$$T = \frac{(4-4.5)^2}{4.5} + \cdots + \frac{(0-0.5)^2}{0.5} = 28/9$$

$$\text{P-value}: P(T > c \mid \text{fit}) \approx 1 - F_{\chi^2_7}(28/9) = 0.874 \rightarrow \text{quite high} \Downarrow \text{a good fit!}$$