

**1**

**A**

Null hypothesis,  $H_0 : \mu = 7$  pounds;

Alternative hypothesis  $H_a = \mu \leq 6$  pounds

**B**

Note that this hypothesis test is left tailed.

If  $H_0$  is true, then  $\bar{x} \sim N(7, \frac{1}{10}) \rightarrow \frac{\bar{x}-7}{1/\sqrt{10}} \sim N(0, 1)$

P-value =  $\Phi((6.2 - 7) * \sqrt{10}) = 0.005706018$

**C**

**D**

No, because if  $H_0$  is true, we can derive the distribution of  $\bar{x}$  without use the CLT.

**E**

Given that I only have the sample variance  $s^2 = 1$ ,

if  $H_0$  is true, then  $\frac{\bar{x}-\mu}{s/\sqrt{n}} \sim t(n-1) \rightarrow \frac{\bar{x}-7}{1/\sqrt{10}} \sim t(9)$

$t = -0.8 * \sqrt{10}, df = 9$

P-value = 0.01612239

**F**

If  $\bar{x} = 6.3$  (given that we know the population variance),

P-value =  $\Phi((6.3 - 7) * \sqrt{10}) = 0.01342835$

Since if the sample mean was less than 6.3 pounds, the decision rule had been to reject the null hypothesis,  $\alpha = 0.01342835$

**G**

If  $\mu = 6.4$

Power =  $1 - P(\text{Type II error}) = 1 - \beta$

$\beta = P(\text{P-value} > \alpha | \mu = 6.4) = P(\bar{x} > 6.3) = P(\frac{\bar{x}-6.4}{1/\sqrt{10}} > \frac{6.3-6.4}{1/\sqrt{10}}) =$

$1 - \Phi(-0.1 * \sqrt{10})$

Power =  $\Phi(-0.1 * \sqrt{10}) = 0.3759148$

## 2

When  $\mu = \mu_0$ ,  $\bar{X} \sim N(\mu_0, \frac{4}{2})$

$$P(\bar{X} > \mu_0 + 1) = P(\frac{\bar{X} - \mu_0}{\sqrt{2}} > \frac{1}{\sqrt{2}}) = 1 - \Phi(1/\sqrt{2}) = 0.2397501$$

$$P(\text{Type 1 error}) = P(\bar{X} > \mu_0 + 1 | \mu = \mu_0) = 0.2397501$$

## 3

### A

$$H_0 : \mu = 3\% \quad H_1 : \mu < 3\%$$

Since the sample size is large enough(200), using the CLT:

$$\frac{\bar{x} - \mu}{s/\sqrt{n}} \sim t(199)$$

$$t = \frac{2.75\% - 3\%}{2\%/\sqrt{200}} = -1.767767$$

$$\text{P-value} = 0.03931615 > 0.01$$

So there is no significant evidence against the null hypothesis.

### B

$$\begin{aligned} \text{power} &= 1 - \beta \\ &= 1 - P(\text{Type 2 error}) \\ &= 1 - P(\text{P-value} > 0.01 | \mu = 2.5\%) \end{aligned}$$

Since  $\frac{\bar{x} - \mu}{s/\sqrt{n}} \sim t(199)$

$$\begin{aligned} P(\text{P-value} > 0.01 | \mu = 2.5\%) &= P(\frac{\bar{x} - 3\%}{s/\sqrt{n}} > -2.345232 | \mu = 2.5\%) \\ &= P(\bar{x} - 3\% > -0.003316659 | \mu = 2.5\%) \\ &= P(\bar{x} - 2.5\% > 0.001683341 | \mu = 2.5\%) \\ &= P(\frac{\bar{x} - 2.5\%}{2\%/\sqrt{200}} > \frac{0.001683341}{2\%/\sqrt{200}} | \mu = 2.5\%) \\ &= 1 - 0.8823271 \end{aligned}$$

$$\text{Power} = 1 - (1 - 0.8823271) = 0.8823271$$

## C

$$P\left(\frac{\bar{x}-3\%}{s/\sqrt{n}} > -2.345232 \mid \mu = 2.5\%\right) = 0.05$$

$$\text{Since } \frac{\bar{x}-\mu}{s/\sqrt{n}} \sim t(199)$$

$$s = 2\%$$

$$P\left(\frac{\bar{x}-2.5\%}{s/\sqrt{n}} > \frac{0.5\%}{s/\sqrt{n}} - 2.345232 \mid \mu = 2.5\%\right) = 0.05$$

$$t_{0.05}(199) = -1.652547$$

$$\frac{0.5\%}{s/\sqrt{n}} - 2.345232 = 1.652547$$

$$n = 256$$

## D

a=the proportion of individuals claiming charitable deductions

X=the number of samples that did not claim any deductions for charitable contributions

$$X \sim B(200, a)$$

Since the sample size is large enough(200), using the CLT:

$$\bar{X} = X/200 \sim N(a, a(1-a)/200)$$

$$H_0 : a = 80\% \quad H_1 : a > 80\%$$

$$\text{P-value} = 1 - \Phi\left(\frac{\bar{X}-a}{\sqrt{\frac{a(1-a)}{200}}}\right) = 1 - \Phi(2.65165) = 0.004004976 < 0.01$$

So there is significant evidence to support that the proportion of individuals claiming charitable deductions is higher in the current year compared to the prior year.

## 4

Assume the mean of difference in systolic blood pressure is  $\mu_0 - \mu_1$ ,

If the sample size is large enough, using CLT:

the mean difference in systolic blood pressure follows  $N(\mu_0 - \mu_1, \frac{16^2}{n} + \frac{16^2}{n})$

$$H_0 : \mu_0 - \mu_1 = 5 \quad H_1 : \mu_0 - \mu_1 < 5$$

$$\text{P-value} = \Phi\left(\frac{\bar{x}-5}{\sqrt{16^2/n+16^2/n}}\right) = 0.05 = \Phi(-1.644854)$$

$$\frac{\bar{x}-5}{\sqrt{16^2/n+16^2/n}} = -1.644854$$

$$P(\text{type 2 error}) = 1 - \text{Power} = 0.1 = \Phi(-1.281552)$$

$$P(\text{P-value} > 0.05 \mid \mu = \mu_0)$$

assume  $\mu_0 = 0$

$$P\left(\frac{\bar{x}-5}{\sqrt{16^2/n+16^2/n}} > -1.644854 \mid \mu_0 = 0\right) = 1 - \Phi(1.281552)$$

$$\frac{5}{\sqrt{16^2/n+16^2/n}} - 1.644854 = 1.281552$$

$n = 176$  (The sample size do large enough to use the CLT)

## 5

### A

$H_0$  : the deck has the right number of hearts

$H_1$  : the deck doesn't have the right number of hearts

n=the total number of hearts

If  $H_0$  is true,  $n \sim B(200, 0.25)$

since the sample size is large enough, using CLT:

$n \sim N(50, 37.5)$

P-value =  $2(1 - \Phi(\frac{64-50}{\sqrt{37.5}})) = 0.02224314 \approx 2.2\%$

### B

False.

Since  $P(A|B) = P(B|A)$  is not always true, P-value=2.2% means given that the null hypothesis is true, there is a 2.2 percent chance of getting at least the same extreme situation that 64 hearts were drawn from the cards.

### C

False.

If the deck is an ordinary deck in the way described above, the chance of getting 64 or more hearts is 1.1 percent.

## 6

### A

Paired or Matched pairs t/Z-test.(One tailed)

Two kinds of experiments have been conducted on the same group of drivers.

We want to find out if the new route is quicker on average than the standard route or not, So it is one-tailed.

### B

Two sample t/Z-test of population mean differences. (One-tailed)

Two kinds of experiments have been conducted on two different group of drivers.

We want to find out if the new route is quicker on average than the standard route or not, So it is one-tailed.

### C

One sample t/Z-test of a population mean.(One-tailed)

## D

Two sample Z-test of population proportion differences.(One-tailed)

Two groups were sampled and the method is comparing the differences of the proportion of people that improved their condition.

We want to find out if the new drug is better than the standard drug, So it is one-tailed.

## E

One sample Z-test of a population proportion.(One-tailed)

The target of the test is to find out if the proportion of cracked tiles is larger than 10%, So we should use one-tailed one sample Z-test of a population proportion.

## 7

### A

Assume that the tar content of this kind of cigarette follows normal distribution  $N(\mu, \sigma^2)$ .

$$\bar{x} = 14.4 \quad s^2 = 0.025 \quad t = \frac{0.4}{\sqrt{0.025}/\sqrt{5}}$$

$$\frac{\bar{x}-\mu}{s/\sqrt{n}} \sim t(4)$$

$$\text{P-value} = 2(1 - 0.9975937) = 0.0048126 < 0.05$$

So for  $\alpha = 0.05$  the null hypothesis  $\mu = 14.0$  will be rejected in favor of the alternative  $\mu \neq 14.0$ .

### B

Yes.

Because without the assumption that the data are drawn from a normal population I can't get the distribution of the estimator I used.

### C

$$\bar{x}_1 = 14.7 \quad s^2 = 0.55 \quad t = \frac{0.7}{\sqrt{0.55}/\sqrt{5}}$$

$$\frac{\bar{x}-\mu}{s/\sqrt{n}} \sim t(4)$$

$$\text{P-value} = 2(1 - 0.9487882) = 0.1024236 > 0.05$$

So the difference is no longer statistically significant.

Because the standard deviation also increased, and the increment of standard deviation cover the effect of increment in mean difference and eventually lead to the reversed result.

## 8

$p_m$  = the number of men in favor of the amendment;

$p_f$  = the number of women in favor of the amendment;

$X_m$  = the proportion of men in favor of the amendment;

$X_f$  = the proportion of women in favor of the amendment;

$X_m \sim B(100, p_m), \quad X_f \sim B(100, p_f)$

Since the two sample sizes are both large enough(100), using the CLT:

$X_m \sim N(100p_m, 100p_m(1 - p_m)) \quad X_f \sim N(100p_f, 100p_f(1 - p_f))$

$(X_m - X_f)/100 \sim N(p_m - p_f, (p_m(1 - p_m) + p_f(1 - p_f))/100)$

Since the confidence interval for this difference (men minus women) was (0.04 0.10).

$$\frac{0.07 - (p_m - p_f)}{\sqrt{(p_m(1 - p_m) + p_f(1 - p_f))/100}} = \pm Z_{\alpha/2}$$

$$\begin{aligned} p_m - p_f &= 0.07 \pm Z_{\alpha/2} * \sqrt{(p_m(1 - p_m) + p_f(1 - p_f))/100} \\ &= 0.07 \pm Z_{\alpha/2} * \sqrt{\frac{0.61 * 0.39 + 0.54 * 0.46}{100}} \\ &= 0.07 \pm Z_{\alpha/2} * 0.06973521 \end{aligned}$$

$$|Z_{\alpha/2}| * 0.06973521 * 2 = 0.10 - 0.04 = 0.06$$

$$Z_{\alpha/2} = -0.4301987 \quad \alpha/2 = 0.3335256 \quad \alpha = 0.6670512$$

The level of confidence:

$$1 - \alpha = 0.3329488$$