9.6

In order to compare the means of two populations, independent random samples of 400 observations are selected from each population, with the following results:

Sample 1	Sample 2
$\bar{x}_1 = 5{,}275$	$\bar{x}_2 = 5{,}240$
$s_1 = 150$	$s_2 = 200$

a. Use a 95% confidence interval to estimate the difference between the population means $\mu_1 - \mu_2$. Interpret the confidence interval.
b. Test the null hypothesis $H_0 : (\mu_1 - \mu_2) = 0$ versus the alternative hypothesis $H_a : (\mu_1 - \mu_2) \neq 0$. Say as much as you can about the p-value of the test, and interpret the result.
c. Suppose the test in part **b** were conducted with the alternative hypothesis $H_a : (\mu_1 - \mu_2) > 0$. How would your answer to part **b** change?
d. Test the null hypothesis $H_0 : (\mu_1 - \mu_2) = 25$ versus the alternative $H_a : (\mu_1 - \mu_2) \neq 25$. Give the p-value, and interpret the result. Compare your answer with that obtained from the test conducted in part **b**.
e. What assumptions are necessary to ensure the validity of the inferential procedures applied in parts **a**–**d**?

a. $CI_{95\%}(\mu_1-\mu_2) = \left(\bar{x}-\bar{y} \mp z_{0.025}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}\right) =$

$n_1 = n_2 = 400$
LARGE SAMPLES

$z_{0.025} = 1.96$

$= \left(5275 - 5240 \mp 1.96\sqrt{\frac{150^2}{400}+\frac{200^2}{400}}\right) =$

$= (35 \mp 24.5)$

b. The rejection region for the test $H_0: \mu_1-\mu_2=0$ vs. $H_a: \mu_1-\mu_2 \neq 0$ is $R=\{|z|> z_{\alpha/2}\}$ for a significance level α, where the test statistic is:

$$z = \frac{\bar{x}-\bar{y}}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}} = \frac{5275-5240}{\sqrt{\frac{150^2}{400}+\frac{200^2}{400}}} = 2.8$$

$\alpha = 0.01 \longrightarrow z_{\alpha/2} = z_{0.005} = 2.575 \rightarrow$ We reject H_0 at the 1% significance level

$\alpha = 0.005 \rightarrow z_{\alpha/2} = z_{0.0025} = 2.81 \rightarrow$ We can't reject H_0 at the 0.5% significance level

$0.005 < \text{p-value} < 0.01$

c. Then the rejection region would be $R=\{z > z_\alpha\}$ with z the same as in b.

$\alpha = 0.005 \rightarrow z_\alpha = 2.575 \rightarrow$ We reject $H_0: \mu_1-\mu_2 \leq 0$ at the 0.5% level

$\alpha = 0.0025 \rightarrow z_\alpha = z_{0.0025} = 2.81 \rightarrow$ We don't reject H_0 at the 0.25% level.

$0.0025 < \text{p-value} < 0.005$

d. For the test $H_0: \mu_1 - \mu_2 = 25$ vs. $H_a: \mu_1 - \mu_2 \neq 25$ the rejection region at the α significance level is $R = \{|z| > z_{\alpha/2}\}$ where the test statistic is

$$z = \frac{\bar{x} - \bar{y} - 25}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{5275 - 5240 - 25}{\sqrt{\frac{150^2}{400} + \frac{200^2}{400}}} = 0{,}8$$

For $\alpha = 0.1 \rightarrow z_{\alpha/2} = z_{0.05} = 1.645 \rightarrow$ We don't reject $H_0: \mu_1 - \mu_2 = 25$ at the 10% significance level.

For $\alpha = 0{,}4 \rightarrow z_{\alpha/2} = z_{0,2} = 0.84 \rightarrow$ We don't reject H_0 at the 40% signif. level

For $\alpha = 0{,}5 \rightarrow z_{\alpha/2} = z_{0,25} = 0{,}675 \rightarrow$ We reject H_0 at the 50% signif. level

$0{,}4 < \text{pvalue} < 0{,}5$

e. We don't need any assumptions, except that X and Y are independent. The CI and the rejection regions are derived using the CLT.