Inference Based on Two Samples

Author: CHEA MAKARA Academy Year:2022-2023

Solution

1). Consider the hypothesis test $H_0: \mu_1 = \mu_2$ against $H_a: \mu_1 \neq \mu_2$ with known variances $\sigma_1 = 10$ and $\sigma_2 = 5$. Suppose that sample sizes $n_1 = 10$ and $n_2 = 15$ and that $\bar{x}_1 = 4.7$ and $\bar{x}_2 = 7.8$. Use $\alpha = 0.05$.

(a). Test the hypotheses and find P-value.
Test $H_0: \mu_1 - \mu_2 = 0$ versus $H_a: \mu_1 - \mu_2 \neq 0$
Test statistic value $z = \dfrac{\bar{x}_1 - \bar{x}_2 - \Delta_0}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} = \dfrac{4.7 - 7.8}{\sqrt{10^2/10 + 5^2/15}} = \dfrac{-3.1}{3.41} = -0.9$
P-value $= 2(1 - \phi(|z|)) = 2(1 - \phi(0.9)) = 0.34$
Since p-value $> \alpha$. Then we decide to not reject H_0 based on the given sample
(b). Explain how the test could be conducted with a confidence interval.
we have $1 - \alpha = P(-z_{\alpha/2} < z < z_{\alpha/2})$
$\iff 1 - \alpha = P(-z_{\alpha/2} < \dfrac{\bar{x}_1 - \bar{x}_2 - \Delta_0}{\sigma} < z_{\alpha/2})$ where $\Delta = \mu_1 - \mu_2$ and $\sigma = \sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}$
$\implies 1 - \alpha = P(-z_{\alpha/2}\sigma + (\bar{x}_1 - \bar{x}_2) > \mu_1 - \mu_2 > (\bar{x}_1 - \bar{x}_2) + z_{\alpha/2}\sigma)$
We obtain $(1 - \alpha)100\%$ for $\mu_1 - \mu_2$ defined by
$CI(\mu_1 - \mu_2) = [(\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2}\sigma]$ where $\sigma = \sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}$

(c). What is the power of the test in part(a) for a true difference in means of 3
we have $\pi(\Delta') = 1 - \beta(\Delta')$, where $\beta(\Delta') = \phi(z_{\alpha/2} - \dfrac{\Delta' - \Delta_0}{\sigma}) - \phi(-z_{\alpha/2} - \dfrac{\Delta' - \Delta_0}{\sigma})$
We have $\Delta' = 3, \Delta_0 = 0, \alpha = 0.05 \implies z_{\alpha/2} = 1.96$

(d). Assuming equal sample sizes, what sample size should be used to obtain $\beta = 0.05$ if there is true difference in means is 3 ? Assume that $\alpha = 0.05$
We have $\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2} = \dfrac{(\Delta' - \Delta_0)^2}{(z_\alpha + z_\beta)^2}$
Suppose $n_1 = n_2$ then we obtain $\dfrac{\sigma_1^2 + \sigma_2^2}{n_1} = \dfrac{(\Delta' - \Delta_0)^2}{(z_\alpha + z_\beta)^2} \implies n_1 = \dfrac{(\sigma_1^2 + \sigma_2^2)(z_\alpha + z_\beta)^2}{(\Delta' - \Delta_0)^2}$
Since $\Delta' = 3, \Delta_0 = 0, \alpha = \beta = 0.05 \implies z_\alpha = z_\beta = z_{0.05} = \phi^{-1}(0.95) = 1.65$
Then $n_1 = \dfrac{(10^2 + 5^2)(1.65 + 1.65)^2}{(3 - 0)^2} = 151.25 \approx 152$
Therefore the sample is $n_1 = n_2 = 152$ ■

---

2). Consider the hypotheses test $H_0: \mu_1 = \mu_2$ against $H_a: \mu_1 > \mu_2$ with known variances $\sigma_1 = 10$ and $\sigma_2 = 5$. Suppose that sample sizes $n_1 : 10$ and $n_2 = 15$ and that $\bar{x}_1 = 24.5$ and $\bar{x}_2 = 21.3$. Use $\alpha = 0.05$
(a). Test the hypotheses and find the P-value.
Proof
We have $H_0: \mu_1 - \mu_2 = 0$ versus $\mu_1 - \mu_2 > 0$
then value of testing statistic $z = \dfrac{\bar{x}_1 - \bar{x}_2 - \Delta_0}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$
Since $\bar{x}_1 = 24.5$ and $\bar{x}_2 = 21.3, \Delta_0 = 0, n_1 = 10, n_2 = 15, \sigma_1 = 10, \sigma_2 = 5$
so $z = \dfrac{24.5 - 21.3 - 0}{\sqrt{10^2/10 + 5^2/15}} = \dfrac{3.2}{\sqrt{11.66}} = 0.93$
P-value is define by $p - value = P(Z > z) = 1 - P(Z < z) = 1 - \phi(z) = 1 - \phi(0.93) = 1 - 0.82 = 0.18$
Since $p - value > \alpha = 0.05$,
Therefore we decide to accept H_0 at level α based on the given sample ■
(b). Explain how the test could be conducted with a confidence interval.
We have $1 - \alpha = P(-z_{\alpha/2} < Z < z_{\alpha/2})$
$\iff 1 - \alpha = P(-z_{\alpha/2} < \dfrac{\bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} < z_{\alpha/2})$
$\iff 1 - \alpha = P(-z_{\alpha/2} \times \sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2} < \bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2) < z_{\alpha/2} \times \sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2})$

$\Longleftrightarrow 1-\alpha = P(\bar{x}-1-\bar{x}_2 - z_{\alpha/2}\times\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2} < \mu_1-\mu_2 < \bar{x}_1-\bar{x}_2+z_{\alpha/2}\times\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2})$
Hence the $100(1-\alpha)\%$ is defined by
$CI(\mu_1-\mu_2) = \left[\bar{x}_1-\bar{x}_2 \pm z_{\alpha/2}\times\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}\right]$
Since $alpha = 0.05 \Longrightarrow z_{0.05} = \phi^{-1}(0.95) = 1.65$
Then $CI = \left[3.2 \pm 1.65\sqrt{11.66}\right] = \left[-2.43, 8.83\right]$ ■
(c). What is the power of the test in part(a) if $\mu$is 2 units greater than μ_2
We have $\pi(\Delta') = 1-\beta(\Delta')$
For testing $\mu_1-\mu_2>0$ then $\beta(\Delta') = \phi(z_\alpha - \frac{\Delta'-\Delta_0}{\sigma})$, where $\sigma = \sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}$
we have $\beta(2) = \phi(1.65 - \frac{2-0}{\sqrt{11.66}}) = \phi(1.06) = 0.855$
then $\phi(2) = 1-\beta(2) = 1-0.855 = 0.145$ ■
(d). Assuming equal sample sizes, what sample size should be used to obtain $\beta = 0.05$ if μ_1 is 2 units greater than μ_2? Assume that $\alpha = 0.05$
We have $\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2} = \frac{(\Delta'-\Delta_0)^2}{(z_\alpha+z_\beta)^2}$
Suppose $n_1 = n_2$ then we obtain $\frac{\sigma_1^2+\sigma_2^2}{n_1} = \frac{(\Delta'-\Delta_0)^2}{(z_\alpha+z_\beta)^2} \Longrightarrow n_1 = \frac{(\sigma_1^2+\sigma_2^2)(z_\alpha+z_\beta)^2}{(\Delta'-\Delta_0)^2}$
Since $\Delta' = 2, \Delta_0 = 0, \alpha = \beta = 0.05 \Longrightarrow z_\alpha = z_\beta = z_{0.05} = \phi^{-1}(0.95) = 1.65$
so $n = \frac{((10^2+5^2)(1.65+1.65)^2)}{(2-0)^2} = 340.3 \approx 341$
Therefore $n = n_1 = n_2 = 341$ ■

---

(3). Two machines are used for filling plastic bottles with a net volume of 16.0 ounces. The fill volume can be assumed normal, with standard deviation $\sigma_1 = 0.020$ and $\sigma_2 = 0.025$ ounces. A member of the equality engineering staff suspects that both machines fill to the same mean net volume, whether or not this volume is 16.0 ounces. A random sample of 10 bottles is taken from the output of each machine

Machine	1	Mchine	2
16.03	16.01	16.02	16.03
16.04	15.96	15.97	16.04
16.05	15.98	15.96	16.03
16.05	16.02	16.01	16.01
16.02	15.99	15.99	16.00

(a). Do you think the engineer is correct ? Use $\alpha = 0.05$ What is the P-value for this test?
Test $H_0: \mu_1-\mu_2 = 0$ vs $H_a: \mu_1-\mu_2 \neq 0$
Test statistic value $z = \frac{\bar{x}_1-\bar{x}_2-\Delta_0}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}}$
We have $\bar{x}_1 = 16.015, \sigma_1 = 0.03$ and $\bar{x}_2 = 16.005, \sigma_2 = 0.025$
Then $z = \frac{16.015-16.005-0}{\sqrt{0.03^2/10+0.025^2/10}} = \frac{0.001}{0.012} = 0.0833$
Find the critical region $C = \{z : |z| \geq z_{\alpha/2}\}$
For significance level $\alpha = 0.05 \Longrightarrow z_{\alpha/2} = z_{0.025} = \phi^{-1}(0.975) = 1.96$
Since static value $z = 0.083 < 1.96$ then $z \neq C$.
Hence we failed to reject H_0 based on the given sample.
Find p-value
we have $p-value = P(|Z| \geq z) = 2(1-\phi(|z|)) = 2(1-\phi(0.0833)) = 2(1-0.798) = 0.404$
Since $P-value > \alpha = 0.05$
Therefore H_0 is not rejected ■
(b). Calculate a 95% confidence interval on the difference in means. Provide a practical interpretation of this interval.
We have $CI = [\bar{x}_1 - \bar{x}_2 \pm z_{\alpha/2}\sigma]$
where $\sigma = \sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}$
then $CI = [16.015-16.005 \pm 1.96\times 0.012] = [-0.013, 0.033]$

We have $0 \in CI$ which means the null hypothesis exists on this interval ■
(c). What is the power of the test in part (a) for a true difference in means 0.04?
We have $\pi(\Delta') = 1 - \beta(\Delta')$
$\Longrightarrow \pi(0.04) = 1 - \beta(0.04)$
where $\beta(\Delta') = phi(z_{\alpha/2} - \frac{\Delta' - \Delta_0}{\sigma}) - \phi(-z_{\alpha/2} - \frac{\Delta' - \Delta_0}{\sigma}) \Longrightarrow \beta(0.04) = \phi(1.96 - \frac{0.04}{0.012}) - \phi(-1.96 - \frac{0.04}{0.012}) = \phi(-1.33) - \phi(-5.26) = -\phi(1.33) + \phi(5.26) = 1 - 0.908 = 0.092$
Then $\pi(0.04) = 1 - 0.092 = 0.908$ ■
(d). Assuming equal sample sizes, what sample size should be used to assure that $\beta = 0.05$ if the true difference in means is 0.04? Assume that $\alpha = 0.05$
We have $\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} = \frac{(\Delta' - \Delta_0)^2}{(z_\alpha + z_\beta)^2}$
Suppose $n_1 = n_2$ then we obtain $\frac{\sigma_1^2 + \sigma_2^2}{n_1} = \frac{(\Delta' - \Delta_0)^2}{(z_\alpha + z_\beta)^2} \Longrightarrow n_1 = \frac{(\sigma_1^2 + \sigma_2^2)(z_\alpha + z_\beta)^2}{(\Delta' - \Delta_0)^2}$
Since $\Delta' = 0.04, \Delta_0 = 0, \alpha = \beta = 0.05 \Longrightarrow z_\alpha = z_\beta = z_{0.05} = \phi^{-1}(0.95) = 1.65$
so $n = \frac{((0.03^2 + 0.025^2)(1.65 + 1.65)^2)}{(0.04 - 0)^2} = 10.37 \approx 11$
Therefore $n = n_1 = n_2 = 11$ ■

---

(4). Two different formulations of an oxygenated motor fuel are being tested to study their road octane members. The variance of the road octane number for formulation 1 is $\sigma_1^2 = 1.5$ and for formulation 2 it is $\sigma_2^2 = 1.2$ Two random samples of size $n_1 = 15$ and $n_2 = 20$ are tested, and the mean road octane numbers observed are $\bar{x}_1 = 89.6$and $\bar{x}_2 = 92.5$. Assume normality.
(a). If formulation 2 produces a higher road octane number then formulation 1 , the manufacturer would like to detect it . Formulate and test an appropriate hypothesis, using $\alpha = 0.05$. What is the P-value?
Test the null hypothesis $H_0 : \mu_1 - \mu_2 = 0$ versus alternative hypotheses$H_a : \mu_1 - \mu_2 < 0$
The statistic value : $z = \frac{\bar{x}_1 - \bar{x}_2 - \Delta_0}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$
We have $\bar{x}_1 = 89.6$ and $\bar{x}_2 = 92.5$ with $n_1 = 15, \sigma_1^2 = 1.5 \& \sigma_2^2 = 1.2, n_2 = 20$
then $z = \frac{89.6 - 92.5}{\sqrt{1.5/15 + 1.2/20}} = \frac{-2.9}{0.4} = -7.25$
Critical Region $RR = \{z : z \leq -z_\alpha\}$ We have $\alpha = 0.05 \Longrightarrow z_{0.05} = \phi^{-1}(0.95) = 1.65$
So , $RR = \{z : z \leq -1.65\}$, Since $z = -7.25 \in RR$. So We can reject the null hypotheses H_0
Find p-value For left side alternative then p-value = $\phi(z) = \phi(-7.25) \approx 0 < \alpha = 0.05$, Then H_0 is also rejected by using p-value method
Therefore the formulation 2 is a higher road octane number than formulate 1 . ■

(b). Explain how the equation in part (a) could be answered with a 95% confidence interval on the difference in mean road octane number.
Find the confidence interval
We have CI=$\left[\bar{x}_1 - \bar{x}_2 \pm z_{\alpha/2}\sigma\right]$
where $z_{\alpha/2} = z_{0.025} = 1.96$ with $\sigma = 0.4$
then $CI = \left[89.6 - 92.5 \pm 1.96\right] = \left[-3.684, -2.116\right]$
We can says $0 \neq CI$ which means the value of $\mu_1 - \mu_2$ is not equal to zero ■

(c). What sample size would be required in each population if we wanted to be 95% confident that the error in estimating the difference in mean road octane number is less that 1 .
We have $n_1 = (\frac{\sigma_1 z_\alpha}{\Delta'})^2 = (\frac{1.65 \times \sqrt{1.5}}{1})^2 = 4.08 \approx 5$
$n_2 = (\frac{1.65 \times \sqrt{1.2}}{1})^2 = 3.26 \approx 4$ ■

---

5). The diameter of steel rods manufactured on two different extrusion machines is being investigated. Two random sample of sizes $n_1 = 15$ and $n_2 = 17$ are selected,and the sample means and sample variance are $\bar{x}_1 = 8.73$ and $s_1^2 = 0.35, \bar{x}_2 = 8.68$ and $s_2^2 = 0.40$m respectively. Assume that $\sigma_1^2 = \sigma_2^2$and that the data are drawn from a normal distribution

(a). Is there evidence to support the claim that the two machines produce rods with different mean diameters? use $\alpha = 0.05$ in arriving at this conclusion. Find the P-value.
Test $H_0 : \mu_1 - \mu_2 = 0$ versus $H_a : \mu_1 - \mu_2 \neq 0$
We use pooled-t test where the testing value of statistic is $t = \dfrac{\bar{x}_1 - \bar{x}_2 - 0}{S_D\sqrt{1/n_1 + 1/n_2}}$
we have $S_D = \sqrt{\dfrac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}}$
Since $\bar{x}_1 = 8.73 \& \bar{x}_2 = 8.68$, $s_1^2 = 0.35$ and $s_2^2 = 0.40$
then $S_D = \sqrt{\dfrac{14(0.35) + (16)0.40}{30}} = 0.61$
we get $T = \dfrac{8.73 - 8.63}{(0.61)\sqrt{1/15 + 1/17}} = \dfrac{0.1}{0.22} = 0.45$
p-value $= 2\min\{P(T \geq t|H_0, P(T \leq t|H_0))\} = 2\min\{P(T \geq 0.45), P(T \leq 0.45)\} = 2\min\{(0.32; 0.67)\} = 2 \times 0.32 = 0.64 > \alpha$ Not reject H_0
Therefore we have not enough evidence to support the claim.
(b). Construct a 95% confidence interval for the difference in mean rod diameter. Interpret this interval.
We have $CI = \left[\bar{x}_1 - \bar{x}_2 \pm t_{\alpha/2,\nu}S\right]$ where
$S = \sqrt{s_1^2/n_1 + s_2^2/n_2} = 0.132$
$\nu = \dfrac{(s_1^2/n_1 + s_2^2/n_2)^2}{\frac{(s_1^2/n_1)^2}{n_1-1} + \frac{(s_2^2/n_2)^2}{n_2-1}} = \dfrac{(0.008 + 0.009)^2}{0.0000045 + 0.000005} = 30$
Recall for pooled t test the $\nu = n_1 + n_2 - 2$
then $t_{\alpha/2,\nu} = t_{0.025,30} = 1.96$
So $CI = \left[0.1 \pm 1.96 \times 0.132\right] = [-0.159, 0.35]$
we have zero is belong to the confidence interval which means the null hypotheses is true ■

---

6). An article in Fire Technology investigated two different foam expending agents that can be used in the nozzles of fire-fighting spray equipment. A random sample of five observations with an aqueous film-forming foam (AFFF) had a sample mean of 4.7 and a standard deviation of 0.6. A random sample of five observations with alcohol-type concentrates(ATC) had a sample mean of 6.9 and a standard deviation 0.8

(a). Can you draw any conclusions about differences in mean foam expansion ? Assume that both populations are well represented by normal distributions with the same standard deviations.
Test $H_0 : mu_1 - \mu_2 = 0$ versus $H_a : \mu_1 - \mu_2 \neq 0$
Since both population represent by normal distribution and have the same standard deviation , we use pooled t test ,
Test statistic value $T = \dfrac{\bar{x}_1 - \bar{x}_2}{S_D\sqrt{1/n_1 + 1/n_2}}$
we have $n_1 = 5, \bar{x}_1 = 4.7, s_1 = 0.6$ and $n_2 = 5, \bar{x}_2 = 6.9, s_2 = 0.8$
so , $S_D = \sqrt{\dfrac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{(n_1+n_2-2)}} = \sqrt{\dfrac{4(0.6)^2 + 4(0.8)^2}{8}} = 0.707$
We obtain $T = \dfrac{4.7 - 6.9}{0.707\sqrt{1/5 + 1/5}} = \dfrac{-2.2}{0.44} = -5$
Find the critical region , where $C = \{t : |t| \geq t_{\alpha/2,\nu}\}$
We have $\nu = n_1 + n_2 - 2 = 8$then $t_{0.05/2,8} = t_{0.025,8} = 2.30$
Since $T = |-5| > 2.30$ then $T \in C$
Therefore we decide to reject H_0 at level $\alpha = 0.05$ based on the given sample. ■

(b). Find a 95 % confidence interval on the difference in mean foam expansion of these two agents.
we have $CI(\mu_1 - \mu_2) = \left[\bar{x}_1 - \bar{x}_2 \pm t_{\alpha/2,\nu}S_D\sqrt{1/n_1 + 1/n_2}\right]$
$\Longrightarrow CI = \left[-2.2 \pm 2.30 \times 0.707 \times 0.63\right] = \left[-3.22, -1.17\right]$ ■

---

7). The deflection temperature under load for two different types of plastic pipe is being investigated. Two random sample of 15 pipe specimens are tested, and the deflection temperatures observed are as follows (in °F):

Type1 :	206	188	205	187	194	193	207	185	189	213	192	210	194	178	205
Type2 :	177	197	206	201	180	176	185	200	197	192	198	188	189	203	192

(a). Construct box plots and normal probability plots for the two samples. Do these plots provide support of the assumptions of normality and equal variances? Write a practical interpretation fir these plots.
(Ans: It's normally and equal variance)
(b). Do the data support the claim that the deflection temperature under load for type 1 pipe exceeds that of type 2? In reaching your conclusion, use $\alpha = 0.05$ Calculate a P-value.
Test $H_0 : \mu_1 - \mu_2 = 0$ versus$H_a : \mu_1 - \mu 2 > 0$
we have test statistic value $T = \dfrac{\bar{x}_1 - \bar{x}_2}{S_D\sqrt{1/n_1 + 1/n_2}}$
Since $n_1 = n_2 = 15, \bar{x}_1 = 196.4, \bar{x}_2 = 192.06$ and $s_1 = 10.47, s_2 = 9.43$
then $S_D = \sqrt{\dfrac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{{}_1 + n_2 - 2}} = \sqrt{\dfrac{14(10.47)^2 + 14(9.43)^2}{28}} = 9.96$
Then $T = \dfrac{196.4 - 192.06}{9.43\sqrt{1/15 + 1/15}} = \dfrac{4.34}{3.44} = 1.26$
Find P-value such that P-value $= P(T \geq -t_{\alpha,\nu})$
since $\alpha = 0.05$ and $\nu = 28$ then we got p-value $= P(T \geq -t_{0.05,28}) = P(T \geq 1.70) = 0.05$
since P-value $= \alpha$ then we decide to reject H_0
Therefore deflection type 1 exceeds type 2 based on given sample. ■

(c). If the mean deflection temperature of type 1 pipe exceeds that type 2 by as much as $5°F$, it is important to detect this difference with probability at least 0.90. Is the choice of $n_1 = n_2 = 15$ adequate? Use $\alpha = 0.05$
find sample size
we have $\Delta' = \mu_1 - \mu_2 = 5$ and $\beta(\Delta') = 0.90 \Longleftrightarrow \beta(5) = 0.90$
so $\beta = 0.90$ and $\alpha = 0.05$
sample size for $n_1 = n_2 = \dfrac{(z_\alpha + z_\beta)^2(S_1^2 + S_2^2)}{\Delta' - \Delta_0} =$

---

8). Two companies manufacture a rubber material intended for use in an automotive application. The part will be subjected to abrasive wear in the field application, so we decide to compare the material produced by each company in a test. Twenty-five samples of material from each company are tested in an abrasion test, and the amount of wear after 1000 cycles in observed. For company 1, the sample mean and standard deviation of wear are $\bar{x}_1 = 20$ milligrams/1000 cycles and $s_1 = 2$ milligrams/1000 cycles, while for company 2 we obtain $\bar{x}_2 = 15$ milligrams/1000 cycles and $s_2 = 8$ milligrams/1000 cycles.

(a). Do the data support the claim that the two companies produce material with different mean wear? Use $\alpha = 0.05$. and assume each population is normally distributed but that their variance are not equal. What is the P-value for this test?
Test the hypotheses $H_0 : \mu_1 = \mu_2$ versus $H_a : \mu_1 \neq \mu_2$
we use t-test in this problem such that
test statistic value $T = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$
since $\bar{x}_1 = 20 \& \bar{x}_2 = 15, n_1 = n_2 = 25$ and $s_1 = 2, s_2 = 8$
we get $T = \dfrac{20 - 15}{\sqrt{2^2/25 + 8^2/25}} = \dfrac{5}{1.66} = 3.01$
Find P-value
For two tailed test we use P-value $= 2\min\{P(T \geq t), P(T \leq t)\}$
(b). Do the data support a claim that the material from company 1 has higher mean wear than the material from company 2? Use the same assumptions as in part (a).
(c). Construct confidence intervals that will address the questions in part (a) and (b) above.

---

9). The thickness of a plastic film (in miles) on a substrate material is thought to be influenced by the

temperature at which the coating is applied. A completely randomized experiment is carried out. Eleven substrates are coated at $125°F$, resulting in a sample mean coating thickness of $\bar{x}_1 = 103.5$ and a random sample standard deviation of $s_1 = 10.2$.Another 13 substrate are coated at $150°F$, for which $\bar{x}_2 = 99$ and $s_2 = 20.1$ are observed. It was originally suspected that raising the process temperature would reduce mean coating thickness.
(a). Do the data support this claim ? Use $\alpha = 0.01$ and assume that the two population standard deviations are not equal.Calculate an approximate P-value for this test.
(b). How could you have answered the question posed regarding the effect of temperature on coating thickness by using a confidence interval? Explain your answer.

---

10).
11). Two different analytical tests can be used to determine the impurity level in steel alloys. Eight specimens are tested using both procedures, and the results are shown in the following tabulations.

Specimen	Test1	Test2
1	1.2	1.4
2	1.3	1.7
3	1.5	1.5
4	1.4	1.3
5	1.7	2.0
6	1.8	2.1
7	1.4	1.7
8	1.3	1.6

(a). Is there sufficient evidence to conclude that tests differ in the mean impurity level, using $\alpha = 0.01$?
(b). Is there evidence to support the claim that Test 1 generates a mean difference 0.1 units lower than Test 2? Use $\alpha = 0.05$
(c). If the mean from Test 1 is 0.1 less than the mean from Test 2, it is important to detect this with probability at least 0.90. Was the use of eight alloys an adequate sample size? If not, how many alloys should have been used?

---

12). For F distribution, find the following:

$$(a).f_{0.25,5,10} \quad (b).f_{0.10,24,9} \quad (c).f_{0.05,8,15}$$
$$(d).f_{0.75,5,10} \quad (e).f_{0.90,24,9} \quad (f).f_{0.95,8,15}$$

---

13). Consider the hypotheses test $H_0 : \sigma_1^2 = \sigma_2^2$ against $H_a : \sigma_1^2 < \sigma_2^2$. Suppose that the sample size are $n_1 = 15$ and $n_2 = 10$, and that $s_1^2 = 23.2$ and $s_2^2 = 28.8$. Use $\alpha = 0.05$. Test the hypotheses and explain how the test could be conducted with a confident interval on σ_1/σ_2
. Test the hypotheses $H_0 : \sigma_1^2 = \sigma_2^2$ versus $H_a : \sigma_1^2 < \sigma_2^2$
For test population variance we use f-test where ,
Test statistic value $f = \dfrac{s_1^2}{s_2^2}$
Since $s_1^2 = 23.2, s_2^2 = 28.8$
then $f = \dfrac{23.2}{28.2} = 0.82$
Find the critical region that defined by
$RR = \{f : f \leq F_{1-\alpha,n_1-1,n_2-1}\}$
We have $\alpha = 0.05, n_1 = 15\&n_2 = 10$
then we get $F_{0.95,14,10} = F_{0.95,14,9} = 0.37 \Longleftrightarrow RR = \{f : f \leq 0.37\}$
We have $f = 0.82 > 0.37$ then $f \neq RR$ then we decide to not reject the null hypotheses
Hence $\sigma_1^2 = \sigma_2^2$
Construct confidence interval
We have confidence interval on σ_1/σ_2
$CI = [\frac{s_1}{s_2}.\frac{1}{\sqrt{F_{\alpha/2,n_1-1,n_2-1}}}, \frac{s_1}{s_2}\sqrt{F_{\alpha/2,n_2-1,n_2-1}}] = [\frac{4.81}{5.36}\frac{1}{\sqrt{F_{0.025,14,9}}}, \frac{4.81}{5.36}\sqrt{F_{0.025,9,14}}]$

$\Longleftrightarrow CI = [0.89/3.79, 0.89 \times 3.20] = [0.23, 2.848]$

14). Consider the hypotheses test $H_0 : \sigma_1^2 = \sigma_2^2$ against $H_a : \sigma_1^2 \neq \sigma_2^2$. Suppose that the sample sizes are $n_1 = 15$ and $n_2 = 15$, and the sample variances are $s_1^2 = 2.3$ and $s_2^2 = 1.9$. Use $\alpha = 0.05$

(a). Test the hypotheses and explain how the test how the test could be conducted with a confidence interval on σ_1/σ_2

(b). What is the power of the test in part (a) if σ_1 is twice as large as σ_2?

(c). Assuming equal sample sizes, what sample size should be used to obtain $\beta = 0.05$ if the σ_2 is half of σ_1?

15). A study was performed to determine whether men women differ in their repeatability in assembling components on printed circuit boards. Random sample of 25 men and 21 women were selected, and each subject assembled the units. The two sample standard deviations of assembly time were $s_{men} = 0.98$ minutes and $s_{women} = 1.02$ minutes.

(a). Is there evidence to support the claim that men and women differ in repeatability for this assembly task? Use $\alpha = 0.02$ and the state any necessary assumptions about the underlying distribution of the data.

Test $H_0 : \sigma_{men}^2 = \sigma_{women}^2$ versus $H_a : \sigma_{men}^2 \neq \sigma_{women}^2$

We use f-test

Test statistic value $f = \dfrac{s_1^2}{s_2^2} = \dfrac{0.98^2}{1.02^2} = 0.92$

The critical region $RR = \{f : f \geq F_{\alpha/2,n_1-1,n_2-1} or. f \leq F_{1-\alpha/2,n_2-1,n_1-1}\}$

$= \{f \geq F_{0.01,24,20} or. f \leq F_{0.99,20,24}\} = \{f : f \geq 4.14 or. f \leq 1.43\}$

Since $f \in RR$ then we decide to reject H_0

Therefore we have enough evidence to reject H_0 at level $\alpha = 0.02$ based on the given sample ■

(b). Find the 98% confidence interval on the ratio of the two variance.Provide an interpretation of the interval.

we have $CI = [\dfrac{S_{men}^2}{S_{women}^2}\dfrac{1}{F_{\alpha/2,n_1-1,n_2-2}}, \dfrac{S_{men}^2}{S_{women}^2}F_{\alpha/2,n_2-1,n_1-1}]$

for 98% then $\alpha = 0.02 \Longrightarrow \alpha/2 = 0.01 \Longleftrightarrow CI = [0.92\dfrac{1}{F_{0.01,24,20}}, 0.92(F_{0.01,20,24})] = [0.22, 3.56]$

16). To measure air pollution in a home, let X and Y equal the amount of suspended particulate matter (in mg/m^3) measured during a 24-hour period in a home in which there is no smoker and a home in which there is a smoker, respectively. We shall test the null hypotheses $H_0 : \sigma_1^2/\sigma_2^2 = 1$ Vs $H_a : \sigma_1^2/\sigma_2^2 > 1$

(a). If a random sample of size $m = 9$ yielded $\bar{x} = 93$ and $S_x = 12.9$ while a random sample of size $n = 11$ yielded $\bar{y} = 132$ and $s_y = 7.1$ define a critical region and give your conclusion if $\alpha = 0.05$

Test hypotheses $H_0 : \sigma_1^2 = \sigma_2^2$ versus $H_a : \sigma_1^2 > \sigma_2^2$

Test hypotheses value $f = \dfrac{s_1^x}{s_2^2} = \dfrac{12.9^2}{7.1^2} = 3.28$

Critical region defined by $C = \{f : f \leq f_{1-\alpha/2,m-1,n-1}\} = \{f : f \leq F_{0.95,8,10}\}$

Hence Critical region $C = \{f : f \leq 0.29\}$

and f is not belong in Critical region

Therefore the null hypotheses not rejected of significance level 0.05 based on given sample. ■

(b). Now test $H_0 : \mu_1 = \mu_2$ against $H_a : \mu_1 < \mu_2$if $\alpha = 0.05$

Test hypotheses $H_0 : \mu_1 = \mu_2$ versus $H_a : \mu_1 < \mu_2$

from previous problem we get $\sigma_1^2 = \sigma_2^2$, then we use pooled t test

Test statistic value $T = \dfrac{\bar{x} - \bar{y}}{S_p\sqrt{1/m + 1/n}}$

Since $\bar{x} = 93 \& \bar{y} = 132$

then $S_p = \sqrt{\dfrac{(m-1)s_1^2 + (n-1)s_2^2}{m+n-1}} = \sqrt{\dfrac{8(12.9)^2 + 10(7.1)^2}{19}} = 9.82$

then $T = \dfrac{93 - 132}{9.82\sqrt{1/9 + 1/11}} = -8.836$

find degree of freedom $\nu = \dfrac{(s_1^2/m + s_2^2/n)^2}{(s_1^2/m)^2/m - 1 + (s_2^2/n_2)/n - 1} = m + n - 2 = 18$

so critical region is defined by $RR = \{t : t < -t_{\alpha,\nu}\} = \{t : t \le -t_{0.05,18}\} = \{t : t \le 1.73\}$
Since $T \in C$ then we decide to reject H_0 based on the two sample ■

---

17). Two different types of injection-molding machines are used to form plastic parts. A part is considered defective if it has excessive shrinkage or is discolored. Two random sample , each of size 300, are selected,and 15 defective parts are found in the sample from machine 1 while 8 defective parts are found in the sample from machine 2.
(a), Is it reasonable to conclude that both machines produce the same fraction of defective parts , using $\alpha = 0.05$? Find the P-value for this test.
Proof
Let p be the number of defective
Test hypotheses $H_0 : p_1 = p_1$ versus $H_a : p_1 \neq p_2$
Test statistic value $z = \dfrac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\hat{q}(1/n_1 + 1/n_2)}}$
we have $\hat{p}_1 = \dfrac{x_1}{n_1} = \dfrac{15}{300} = 0.05$ and $\hat{p}_2 = \dfrac{x_2}{n_2} = \dfrac{8}{300} = 0.026$
and $\hat{p} = \dfrac{x_1 + x_2}{n_1 + n_2} = \dfrac{23}{600} = 0.048$ then $\hat{q} = 1 - \hat{p} = 1 - 0.048 = 0.952$
we get $z = \dfrac{0.05 - 0.01}{\sqrt{0.048(0.952)(1/300 + 1/300)}} = \dfrac{0.04}{0.0174} = 0.229$
Find p-value , where $p - value = 2(1 - \phi|z|) = 2(1 - \phi(|0.229|)) = 0.824 < \alpha = 0.05$ then H_O is not rejected
Therefore both machines produce the same defective parts. ■
(b). Construct a 95% confidence interval on the different in the two fraction defective.
We have $CI = [\hat{p}_1 - \hat{p}_2 \pm z_{\alpha/2}\sqrt{\dfrac{\hat{p}_1\hat{q}_1}{n_1} + \dfrac{\hat{p}_2\hat{q}_2}{n_2}}] = [0.04 \pm z_{0.025}\sqrt{0.05(0.95)/300 + 0.026(0.0.974)/300}]$
$= [0.04 \pm 1.96 \times 0.015] = [0.0106, 0.0694]$ ■
(c).Suppose that $p_1 = 0.05$ and $p_2 = 0.01$ with the sample sizes given here, what is the power of the test for this two-sided alternative?
we have $\phi(p_1, p_2) = 1 - \beta(p_1, p_2)$ where
$\beta(p_1, p_2) = \phi(\dfrac{z_{\alpha/2}\sqrt{\bar{p}\bar{q}(1/n_1 + 1/n_2)} - (p_1 - p - 2)}{\sigma}) - \phi(\dfrac{-z_{\alpha/2}\sqrt{\bar{p}\bar{q}(1/n_1 + 1/n_2)} - (p_1 - p_2)}{\sigma})$
since $\sigma = 0.015$ and $\bar{p} = \dfrac{n_1p_1 + n_2p_2}{n_1 + n_2} = \dfrac{300 \times 0.05 + 300(0.026)}{300 + 300} = 0.038$ and $\bar{q} = \dfrac{n_1q_1 + n_2q_2}{n_1 + n_2} = \dfrac{0.95 + 0.974}{2} = 0.962$
then $\beta = \phi(\dfrac{1.96 \times 0.015}{0.015}) - \phi(\dfrac{-1.96 \times 0.015}{0.015}) = \phi(1.96) - \phi(-1.96) = 2\phi(1.96) - 1 = 0.9$
Therefore $\phi(p_1, p_2) = 1 - 0.9 = 0.1$ ■
(d). Suppose that $p_1 = 0.05$ and $p_2 = 0.01$. Determine the sample size needed to detect this difference with a probability of at least 0.9
we have $\beta = 1 - \phi = 1 - 0.9 = 0.1$
sample size $n = \dfrac{\left[z_{\alpha/2}\sqrt{((p_1 + p_2) + (q_1 + q_2))/2} + z_\beta\sqrt{p_1q_1 + p_2q_2}\right]^2}{d^2} = \dfrac{(1.96 \times 0.96 + 1.29 \times 0.96)^2}{0.04}$
$= 243.36 \approx 244$ ■

---

18). Recent incidents of food contamination have caused great concern among consumers. The article "How Safe Is That chicken? "(Consumer Report, Jan.2010:19-23) reported that 35 of 80 randomly selected Perdue brand broilers tested positively for either campylobacter or salmonella (or both), the leading bacterial cause of food-bone disease, whereas 66 of 80 Tyson brand broilers tested positive.
(a). Does it appear that the true proportion of non-contaminated Perdue broilers differs from that for the Tyson brand? Carry out a test of hypotheses using a significance level 0.01 by obtaining a P-value?
(b). If the true proportions of non-contaminated chickens for the Perdue and Tyson brands are 0.50 anf 0.25 , respectively, how likely is it that the null hypotheses of equal proportions will be rejected when a 0.01 significance level is used and the sample sizes are both 80?