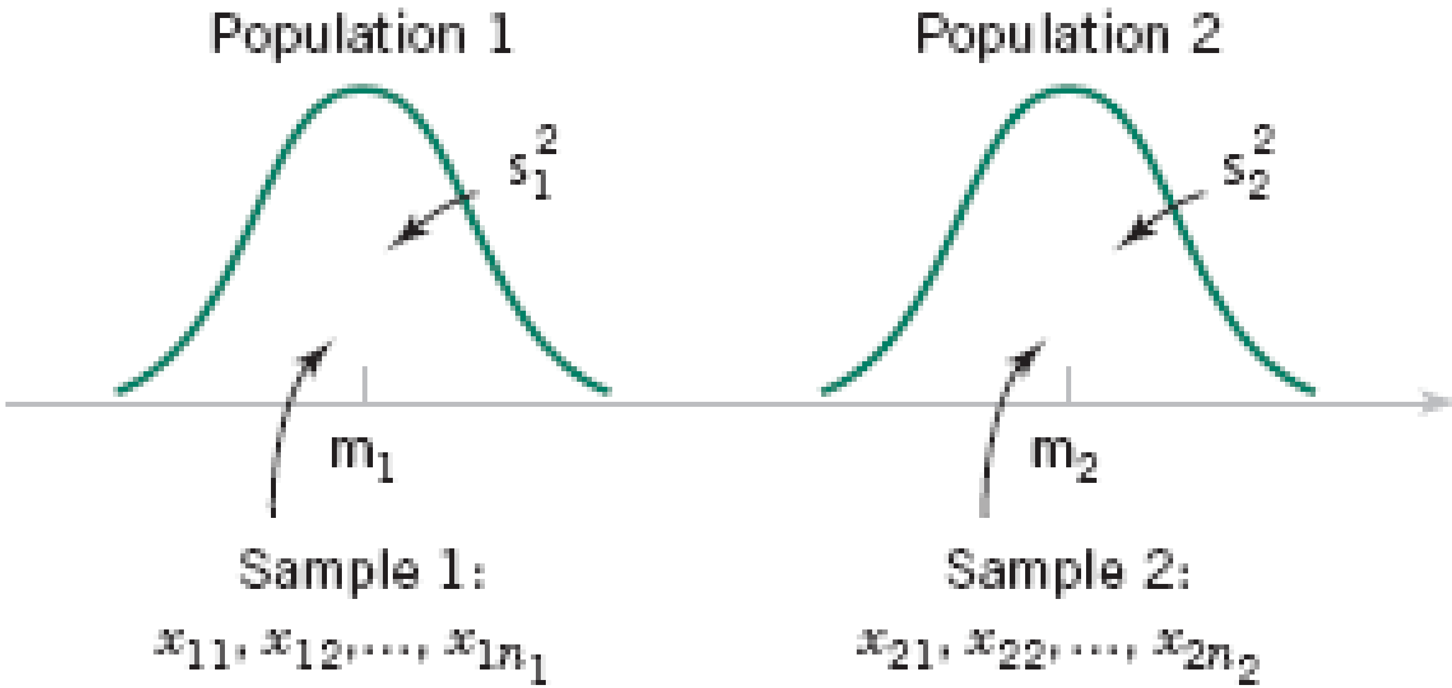
CH.10 Statistical Inference for Two Samples

- Inference on the difference in means of two normal distributions, variances known
- Inference on the difference in means of two normal distributions, variances unknown
- Paired t-test
- Inference on the variances of two normal distributions
- Inference on two population proportions

10-1 Introduction



Two independent populations.

10-1 Inference for a Difference in Means of Two Normal Distributions, Variances Known

Assumptions

1. $X_{11}, X_{12}, \ldots, X_{1n_1}$ is a random sample from population 1.
2. $X_{21}, X_{22}, \ldots, X_{2n_2}$ is a random sample from population 2.
3. The two populations represented by X_1 and X_2 are independent.
4. Both populations are normal.

$$E(\overline{X}_1 - \overline{X}_2) = E(\overline{X}_1) - E(\overline{X}_2) = \mu_1 - \mu_2$$

$$V(\overline{X}_1 - \overline{X}_2) = V(\overline{X}_1) + V(\overline{X}_2) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

10-1 Inference for a Difference in Means of Two Normal Distributions, Variances Known

The quantity

$$Z = \frac{\overline{X}_1 - \overline{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \tag{10-1}$$

has a $N(0, 1)$ distribution.

10-1 Inference for a Difference in Means of Two Normal Distributions, Variances Known

10-1.1 Hypothesis Tests for a Difference in Means, Variances Known

Null hypothesis: $H_0\colon \mu_1 - \mu_2 = \Delta_0$

Test statistic: $$Z_0 = \frac{\overline{X}_1 - \overline{X}_2 - \Delta_0}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \tag{10-2}$$

Alternative Hypotheses	Rejection Criterion
$H_1\colon \mu_1 - \mu_2 \neq \Delta_0$	$z_0 > z_{\alpha/2}$ or $z_0 < -z_{\alpha/2}$
$H_1\colon \mu_1 - \mu_2 > \Delta_0$	$z_0 > z_\alpha$
$H_1\colon \mu_1 - \mu_2 < \Delta_0$	$z_0 < -z_\alpha$

10-1 Inference for a Difference in Means of Two Normal Distributions, Variances Known

Example 10-1

A product developer is interested in reducing the drying time of a primer paint. Two formulations of the paint are tested; formulation 1 is the standard chemistry, and formulation 2 has a new drying ingredient that should reduce the drying time. From experience, it is known that the standard deviation of drying time is 8 minutes, and this inherent variability should be unaffected by the addition of the new ingredient. Ten specimens are painted with formulation 1, and another 10 specimens are painted with formulation 2; the 20 specimens are painted in random order. The two sample average drying times are $\bar{x}_1 = 121$ minutes and $\bar{x}_2 = 112$ minutes, respectively. What conclusions can the product developer draw about the effectiveness of the new ingredient, using $\alpha = 0.05$?

We apply the eight-step procedure to this problem as follows:

1. The quantity of interest is the difference in mean drying times, $\mu_1 - \mu_2$, and $\Delta_0 = 0$.
2. H_0: $\mu_1 - \mu_2 = 0$, or H_0: $\mu_1 = \mu_2$.
3. H_1: $\mu_1 > \mu_2$. We want to reject H_0 if the new ingredient reduces mean drying time.
4. $\alpha = 0.05$

10-1 Inference for a Difference in Means of Two Normal Distributions, Variances Known

Example 10-1

5. The test statistic is

$$z_0 = \frac{\bar{x}_1 - \bar{x}_2 - 0}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$$

where $\sigma_1^2 = \sigma_2^2 = (8)^2 = 64$ and $n_1 = n_2 = 10$.

6. Reject H_0: $\mu_1 = \mu_2$ if $z_0 > 1.645 = z_{0.05}$.

7. Computations: Since $\bar{x}_1 = 121$ minutes and $\bar{x}_2 = 112$ minutes, the test statistic is

$$\boxed{z_0 = \frac{121 - 112}{\sqrt{\dfrac{(8)^2}{10} + \dfrac{(8)^2}{10}}} = 2.52}$$

10-1 Inference for a Difference in Means of Two Normal Distributions, Variances Known

Example 10-1

8. Conclusion: Since $z_0 = 2.52 > 1.645$, we reject H_0: $\mu_1 = \mu_2$ at the $\alpha = 0.05$ level and conclude that adding the new ingredient to the paint significantly reduces the drying time. Alternatively, we can find the P-value for this test as

$$P\text{-value} = 1 - \Phi(2.52) = 0.0059$$

Therefore, H_0: $\mu_1 = \mu_2$ would be rejected at any significance level $\alpha \geq 0.0059$.

10-1 Inference for a Difference in Means of Two Normal Distributions, Variances Known

10-1.2 Type II Error and Choice of Sample Size

Use of Operating Characteristic Curves

Two-sided or one-sided alternative:

$$d = \frac{|\mu_1 - \mu_2 - \Delta_0|}{\sqrt{\sigma_1^2 + \sigma_2^2}} = \frac{|\Delta - \Delta_0|}{\sqrt{\sigma_1^2 + \sigma_2^2}}$$

Choose equal sample sizes, $n=n_1=n_2$

If not possible: $n_1 \neq n_2$, compute equivalent value of n as follows

$$n = \frac{\sigma_1^2 + \sigma_2^2}{\sigma_1^2 / n_1 + \sigma_2^2 / n_2}$$

make adjustment to n_1 and n_2 until the specified β is reached.

10-1 Inference for a Difference in Means of Two Normal Distributions, Variances Known

10-1.2 Type II Error and Choice of Sample Size

Sample Size Formulas

Two-sided alternative:

For the two-sided alternative hypothesis with significance level α, the sample size $n_1 = n_2 = n$ required to detect a true difference in means of Δ with power at least $1 - \beta$ is

$$n \simeq \frac{(z_{\alpha/2} + z_{\beta})^2(\sigma_1^2 + \sigma_2^2)}{(\Delta - \Delta_0)^2} \qquad (10\text{-}5)$$

10-1 Inference for a Difference in Means of Two Normal Distributions, Variances Known

10-1.2 Type II Error and Choice of Sample Size

Sample Size Formulas

One-sided alternative:

For a one-sided alternative hypothesis with significance level α, the sample size $n_1 = n_2 = n$ required to detect a true difference in means of $\Delta(\neq\Delta_0)$ with power at least $1 - \beta$ is

$$n = \frac{(z_\alpha + z_\beta)^2(\sigma_1^2 + \sigma_2^2)}{(\Delta - \Delta_0)^2} \qquad (10\text{-}6)$$

10-1 Inference for a Difference in Means of Two Normal Distributions, Variances Known

Example 10-3

To illustrate the use of these sample size equations, consider the situation described in Example 10-1, and suppose that if the true difference in drying times is as much as 10 minutes, we want to detect this with probability at least 0.90. Under the null hypothesis, $\Delta_0 = 0$. We have a one-sided alternative hypothesis with $\Delta = 10$, $\alpha = 0.05$ (so $z_\alpha = z_{0.05} = 1.645$), and since the power is 0.9, $\beta = 0.10$ (so $z_\beta = z_{0.10} = 1.28$). Therefore we may find the required sample size from Equation 10-6 as follows:

$$n = \frac{(z_\alpha + z_\beta)^2(\sigma_1^2 + \sigma_2^2)}{(\Delta - \Delta_0)^2} = \frac{(1.645 + 1.28)^2[(8)^2 + (8)^2]}{(10 - 0)^2} = 11$$

This is exactly the same as the result obtained from using the O.C. curves.

$$d = \frac{|\mu_1 - \mu_2 - 0|}{\sqrt{\sigma_1^2 + \sigma_2^2}} = \frac{10}{\sqrt{8^2 + 8^2}} = 0.88$$

ALSO,
with β=0.1, d=0.88, α=0.05
from Appendix Chart VIIc
$n=n_1=n_2\approx 11$

10-1 Inference for a Difference in Means of Two Normal Distributions, Variances Known

10-1.3 Confidence Interval on a Difference in Means, Variances Known

Definition

If $\bar{x}_1$ and $\bar{x}_2$ are the means of independent random samples of sizes n_1 and n_2 from two independent normal populations with known variances σ_1^2 and σ_2^2, respectively, a **$100(1-\alpha)\%$ confidence interval for $\mu_1 - \mu_2$** is

$$\bar{x}_1 - \bar{x}_2 - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \le \mu_1 - \mu_2 \le \bar{x}_1 - \bar{x}_2 + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \qquad (10\text{-}7)$$

where $z_{\alpha/2}$ is the upper $\alpha/2$ percentage point of the standard normal distribution.

10-1 Inference for a Difference in Means of Two Normal Distributions, Variances Known

Example 10-4

Tensile strength tests were performed on two different grades of aluminum spars used in manufacturing the wing of a commercial transport aircraft. From past experience with the spar manufacturing process and the testing procedure, the standard deviations of tensile strengths are assumed to be known. The data obtained are as follows: $n_1 = 10$, $\bar{x}_1 = 87.6$, $\sigma_1 = 1$, $n_2 = 12$, $\bar{x}_2 = 74.5$, and $\sigma_2 = 1.5$. If μ_1 and μ_2 denote the true mean tensile strengths for the two grades of spars, we may find a 90% confidence interval on the difference in mean strength $\mu_1 - \mu_2$ as follows:

$$\bar{x}_1 - \bar{x}_2 - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \leq \mu_1 - \mu_2 \leq \bar{x}_1 - \bar{x}_2 + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$87.6 - 74.5 - 1.645\sqrt{\frac{(1)^2}{10} + \frac{(1.5)^3}{12}} \leq \mu_1 - \mu_2 \leq 87.6 - 74.5 + 1.645\sqrt{\frac{(1^2)}{10} + \frac{(1.5)^2}{12}}$$

10-1 Inference for a Difference in Means of Two Normal Distributions, Variances Known

Example 10-4

Therefore, the 90% confidence interval on the difference in mean tensile strength (in kilograms per square millimeter) is

$$12.22 \leq \mu_1 - \mu_2 \leq 13.98 \text{ (in kilograms per square millimeter)}$$

Notice that the confidence interval does not include zero, implying that the mean strength of aluminum grade 1 (μ_1) exceeds the mean strength of aluminum grade 2 (μ_2). In fact, we can state that we are 90% confident that the mean tensile strength of aluminum grade 1 exceeds that of aluminum grade 2 by between 12.22 and 13.98 kilograms per square millimeter.

10-1 Inference for a Difference in Means of Two Normal Distributions, Variances Known

Choice of Sample Size

$$n = \left(\frac{z_{\alpha/2}}{E}\right)^2 (\sigma_1^2 + \sigma_2^2) \qquad (10\text{-}8)$$

The required sample size so that the error in estimating $\mu_1 - \mu_2$ by $\bar{x}_1 - \bar{x}_2$ will be less than E at $100(1-\alpha)\%$ confidence.

10-1 Inference for a Difference in Means of Two Normal Distributions, Variances Known

One-Sided Confidence Bounds

Upper Confidence Bound

$$\mu_1 - \mu_2 \le \bar{x}_1 - \bar{x}_2 + z_\alpha \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \qquad (10\text{-}9)$$

Lower Confidence Bound

$$\bar{x}_1 - \bar{x}_2 - z_\alpha \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \le \mu_1 - \mu_2 \qquad (10\text{-}10)$$

10-2 Inference for a Difference in Means of Two Normal Distributions, Variances Unknown

10-2.1 Hypotheses Tests for a Difference in Means, Variances Unknown

Case 1: $\sigma_1^2 = \sigma_2^2 = \sigma^2$

We wish to test:

$$H_0: \mu_1 - \mu_2 = \Delta_0$$
$$H_1: \mu_1 - \mu_2 \neq \Delta_0$$

10-2 Inference for a Difference in Means of Two Normal Distributions, Variances Unknown

10-2.1 Hypotheses Tests for a Difference in Means, Variances Unknown

Case 1: $\sigma_1^2 = \sigma_2^2 = \sigma^2$

combine S_1^2 and S_2^2 to form an estimator of σ^2

The pooled estimator of σ^2:

The pooled estimator of σ^2, denoted by S_p^2, is defined by

$$\boxed{S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}} \tag{10-12}$$

10-2 Inference for a Difference in Means of Two Normal Distributions, Variances Unknown

10-2.1 Hypotheses Tests for a Difference in Means, Variances Unknown

Case 1: $\sigma_1^2 = \sigma_2^2 = \sigma^2$

Given the assumptions of this section, the quantity

$$T = \frac{\overline{X}_1 - \overline{X}_2 - (\mu_1 - \mu_2)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \tag{10-13}$$

has a t distribution with $\boxed{n_1 + n_2 - 2}$ degrees of freedom.

10-2 Inference for a Difference in Means of Two Normal Distributions, Variances Unknown

Definition: The Two-Sample or Pooled *t*-Test

Null hypothesis: $H_0\colon \mu_1 - \mu_2 = \Delta_0$

Test statistic:
$$T_0 = \frac{\overline{X}_1 - \overline{X}_2 - \Delta_0}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \tag{10-14}$$

Alternative Hypothesis	Rejection Criterion
$H_1\colon \mu_1 - \mu_2 \neq \Delta_0$	$t_0 > t_{\alpha/2, n_1+n_2-2}$ or $t_0 < -t_{\alpha/2, n_1+n_2-2}$
$H_1\colon \mu_1 - \mu_2 > \Delta_0$	$t_0 > t_{\alpha, n_1+n_2-2}$
$H_1\colon \mu_1 - \mu_2 < \Delta_0$	$t_0 < -t_{\alpha, n_1+n_2-2}$

10-2 Inference for a Difference in Means of Two Normal Distributions, Variances Unknown

Example 10-5

Two catalysts are being analyzed to determine how they affect the mean yield of a chemical process. Specifically, catalyst 1 is currently in use, but catalyst 2 is acceptable. Since catalyst 2 is cheaper, it should be adopted, providing it does not change the process yield. A test is run in the pilot plant and results in the data shown in Table 10-1. Is there any difference between the mean yields? Use $\alpha = 0.05$, and assume equal variances.

10-2 Inference for a Difference in Means of Two Normal Distributions, Variances Unknown

Example 10-5

Table 10-1 Catalyst Yield Data, Example 10-5

Observation Number	Catalyst 1	Catalyst 2
1	91.50	89.19
2	94.18	90.95
3	92.18	90.46
4	95.39	93.21
5	91.79	97.19
6	89.07	97.04
7	94.72	91.07
8	89.21	92.75
	$\bar{x}_1 = 92.255$	$\bar{x}_2 = 92.733$
	$s_1 = 2.39$	$s_2 = 2.98$

10-2 Inference for a Difference in Means of Two Normal Distributions, Variances Unknown

Example 10-5

The solution using the eight-step hypothesis-testing procedure is as follows:

1. The parameters of interest are μ_1 and μ_2, the mean process yield using catalysts 1 and 2, respectively, and we want to know if $\mu_1 - \mu_2 = 0$.
2. $\boxed{H_0\colon \mu_1 - \mu_2 = 0}$ or $H_0\colon \mu_1 = \mu_2$
3. $\boxed{H_1\colon \mu_1 \neq \mu_2}$
4. $\alpha = 0.05$
5. The test statistic is

$$t_0 = \frac{\bar{x}_1 - \bar{x}_2 - 0}{s_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$$

6. $\boxed{\text{Reject } H_0 \text{ if } t_0 > t_{0.025,14} = 2.145 \text{ or if } t_0 < -t_{0.025,14} = -2.145.}$

10-2 Inference for a Difference in Means of Two Normal Distributions, Variances Unknown

Example 10-5

7. Computations: From Table 10-1 we have $\bar{x}_1 = 92.255$, $s_1 = 2.39$, $n_1 = 8$, $\bar{x}_2 = 92.733$, $s_2 = 2.98$, and $n_2 = 8$. Therefore

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = \frac{(7)(2.39)^2 + 7(2.98)^2}{8 + 8 - 2} = 7.30$$

$$s_p = \sqrt{7.30} = 2.70$$

and

$$\boxed{t_0 = \frac{\bar{x}_1 - \bar{x}_2}{2.70\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{92.255 - 92.733}{2.70\sqrt{\frac{1}{8} + \frac{1}{8}}} = -0.35}$$

8. Conclusions: Since $-2.145 < t_0 = -0.35 < 2.145$, the null hypothesis cannot be rejected. That is, at the 0.05 level of significance, we do not have strong evidence to conclude that catalyst 2 results in a mean yield that differs from the mean yield when catalyst 1 is used.

10-2 Inference for a Difference in Means of Two Normal Distributions, Variances Unknown

Minitab Output for Example 10-5

Minitab Computations

Two-Sample T-Test and CI: Cat 1, Cat 2

Two-sample T for Cat 1 vs Cat 2

	N	Mean	StDev	SE Mean
Cat 1	8	92.26	2.39	0.84
Cat 2	8	92.73	2.99	1.1

Difference = mu Cat 1 − mu Cat 2

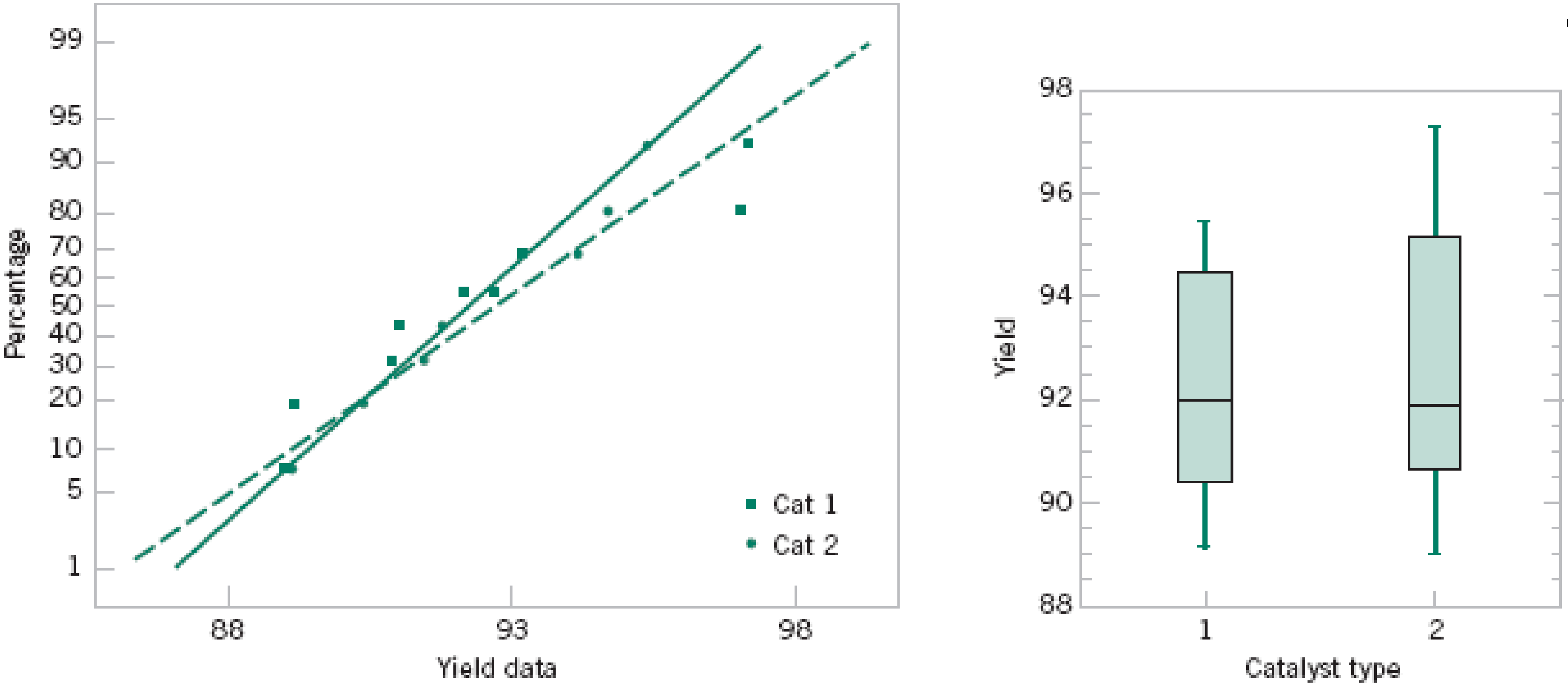
Estimate for difference: −0.48

95% CI for difference: (−3.37, 2.42)

T-Test of difference = 0 (vs not =): T-Value = −0.35 P-Value = 0.730 DF = 14

Both use Pooled StDev = 2.70

10-2 Inference for a Difference in Means of Two Normal Distributions, Variances Unknown



(a)

(b)

Normal probability plot: 1) *normality assumption can be made.*
2) similar slopes indicate assumption of equal variances

Comparative box plot: *no obvious difference in the two catalysts. However, catalyst 2 has slightly greater sample variability.*

10-2 Inference for a Difference in Means of Two Normal Distributions, Variances Unknown

10-2.1 Hypotheses Tests for a Difference in Means, Variances Unknown

Case 2: $\sigma_1^2 \neq \sigma_2^2$

$$T_0^* = \frac{\overline{X}_1 - \overline{X}_2 - \Delta_0}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}} \tag{10-15}$$

is distributed approximately as *t* distribution with degrees of freedom ν given by

10-2 Inference for a Difference in Means of Two Normal Distributions, Variances Unknown

10-2.1 Hypotheses Tests for a Difference in Means, Variances Unknown

Case 2: $\sigma_1^2 \neq \sigma_2^2$

$$v = \frac{\left(\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}\right)^2}{\dfrac{(S_1^2/n_1)^2}{n_1 - 1} + \dfrac{(S_2^2/n_2)^2}{n_2 - 1}} \qquad (10\text{-}16)$$

10-2 Inference for a Difference in Means of Two Normal Distributions, Variances Unknown

Example 10-6

Arsenic concentration in public drinking water supplies is a potential health risk. An article in the *Arizona Republic* (Sunday, May 27, 2001) reported drinking water arsenic concentrations in parts per billion (ppb) for 10 methropolitan Phoenix communities and 10 communities in rural Arizona. The data follow:

Metro Phoenix ($\bar{x}_1 = 12.5, s_1 = 7.63$)	Rural Arizona ($\bar{x}_2 = 27.5, s_2 = 15.3$)
Phoenix, 3	Rimrock, 48
Chandler, 7	Goodyear, 44
Gilbert, 25	New River, 40
Glendale, 10	Apachie Junction, 38
Mesa, 15	Buckeye, 33
Paradise Valley, 6	Nogales, 21
Peoria, 12	Black Canyon City, 20
Scottsdale, 25	Sedona, 12
Tempe, 15	Payson, 1
Sun City, 7	Casa Grande, 18

10-2 Inference for a Difference in Means of Two Normal Distributions, Variances Unknown

Example 10-6 (Continued)

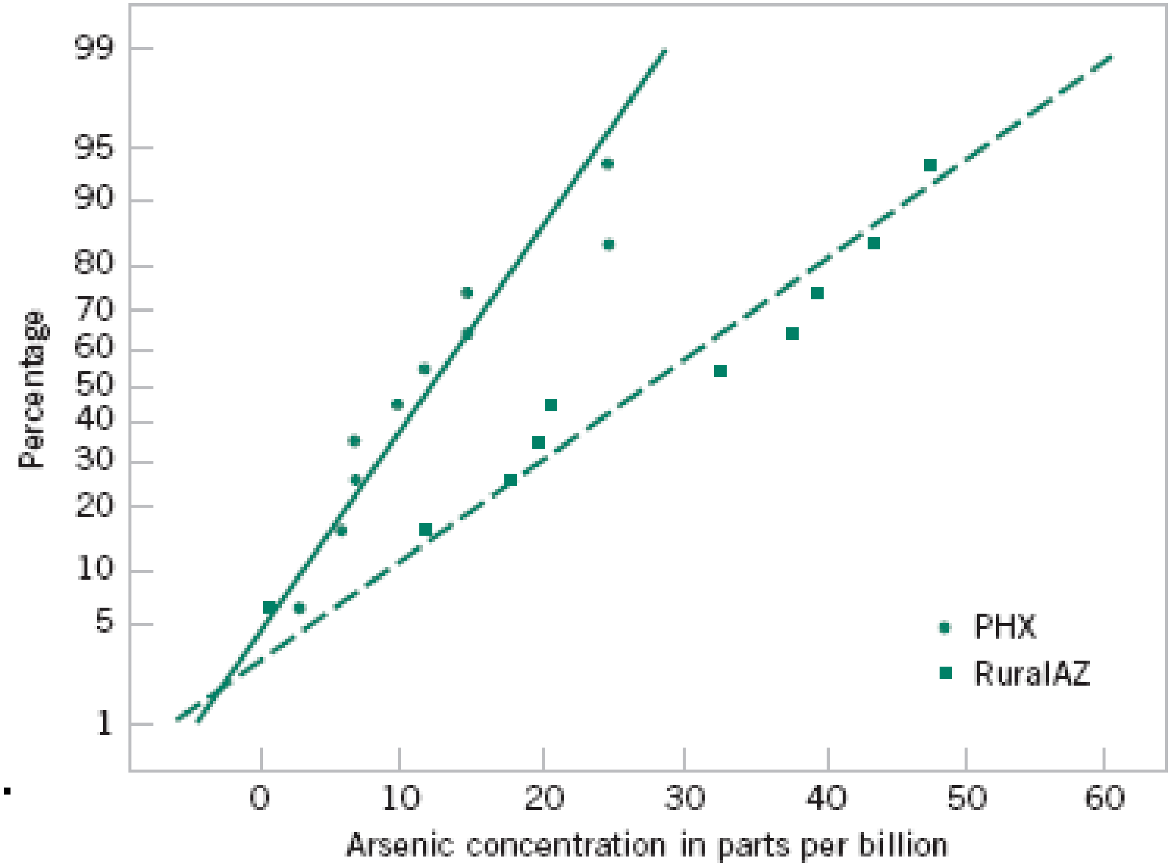


Figure 10-2 Normal probability plot of the arsenic concentration data from Example 10-6.

10-2 Inference for a Difference in Means of Two Normal Distributions, Variances Unknown

Example 10-6 (Continued)

We wish to determine it there is any difference in mean arsenic concentrations between metropolitan Phoenix communities and communities in rural Arizona. Figure 10-3 shows a normal probability plot for the two samples of arsenic concentration. The assumption of normality appears quite reasonable, but since the slopes of the two straight lines are very different, it is unlikely that the population variances are the same.

Applying the eight-step procedure gives the following:

1. The parameters of interest are the mean arsenic concentrations for the two geographic regions, say, μ_1 and μ_2, and we are interested in determining whether $\mu_1 - \mu_2 = 0$.
2. H_0: $\mu_1 - \mu_2 = 0$, or H_0: $\mu_1 = \mu_2$
3. H_1: $\mu_1 \neq \mu_2$
4. $\alpha = 0.05$ (say)

10-2 Inference for a Difference in Means of Two Normal Distributions, Variances Unknown

Example 10-6 (Continued)

5. The test statistic is

$$t_0^* = \frac{\bar{x}_1 - \bar{x}_2 - 0}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$$

6. The degrees of freedom on t_0^* are found from Equation 10-16 as

$$v = \frac{\left(\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}\right)^2}{\dfrac{(s_1^2/n_1)^2}{n_1 - 1} + \dfrac{(s_2^2/n_2)^2}{n_2 - 1}} = \frac{\left[\dfrac{(7.63)^2}{10} + \dfrac{(15.3)^2}{10}\right]^2}{\dfrac{[(7.63)^2/10]^2}{9} + \dfrac{[(15.3)^2/10]^2}{9}} = 13.2 \simeq 13$$

Therefore, using $\alpha = 0.05$, we would reject $H_0\colon \mu_1 = \mu_2$ if $t_0^* > t_{0.025,13} = 2.160$ or if $t_0^* < -t_{0.025,13} = -2.160$

10-2 Inference for a Difference in Means of Two Normal Distributions, Variances Unknown

Example 10-6 (Continued)

7. Computations: Using the sample data we find

$$t_0^* = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{12.5 - 27.5}{\sqrt{\frac{(7.63)^2}{10} + \frac{(15.3)^2}{10}}} = -2.77$$

8. Conclusions: Because $t_0^* = -2.77 < t_{0.025,13} = -2.160$, we reject the null hypothesis. Therefore, there is evidence to conclude that mean arsenic concentration in the drinking water in rural Arizona is different from the mean arsenic concentration in metropolitan Phoenix drinking water. Furthermore, the mean arsenic concentration is higher in rural Arizona communities. The P-value for this test is approximately $P = 0.016$.

10-2 Inference for a Difference in Means of Two Normal Distributions, Variances Unknown

10-2.2 Type II Error and Choice of Sample Size

Case 1: $\sigma_1^2 = \sigma_2^2 = \sigma^2$

Use operating characteristic curves in Appendix Charts VIIe, VIIf, VIIg, VIIh.

$$d = \frac{|\Delta - \Delta_0|}{2\sigma}$$

$$n^* = 2n - 1$$

Case 2: $\sigma_1^2 \neq \sigma_2^2$

No operating characteristic curves are available.

10-2 Inference for a Difference in Means of Two Normal Distributions, Variances Unknown

10-2.2 Type II Error and Choice of Sample Size

Example 10-7

Consider the catalyst experiment in Example 10-5. Suppose that, if catalyst 2 produces a mean yield that differs from the mean yield of catalyst 1 by 4.0%, we would like to reject the null hypothesis with probability at least 0.85. What sample size is required?

Using $s_p = 2.70$ as a rough estimate of the common standard deviation σ, we have $d = |\Delta|/2\sigma = |4.0|/[(2)(2.70)] = 0.74$. From Appendix Chart VII*e* with $d = 0.74$ and $\beta = 0.15$, we find $n^* = 20$, approximately. Therefore, since $n^* = 2n - 1$,

$$n = \frac{n^* + 1}{2} = \frac{20 + 1}{2} = 10.5 \simeq 11(\text{say})$$

and we would use sample sizes of $n_1 = n_2 = n = 11$.

10-2 Inference for a Difference in Means of Two Normal Distributions, Variances Unknown

Minitab Output for Example 10-7

Minitab Computations

Power and Sample Size

2-Sample *t* Test
Testing mean 1 = mean 2 (versus not =)
Calculating power for mean 1 = mean 2 + difference
Alpha = 0.05 Sigma = 2.7

Difference	Sample Size	Target Power	Actual Power
4	10	0.8500	0.8793

10-2 Inference for a Difference in Means of Two Normal Distributions, Variances Unknown

10-2.3 Confidence Interval on the Difference in Means, Variance Unknown

Case 1: $\sigma_1^2 = \sigma_2^2 = \sigma^2$

If $\bar{x}_1, \bar{x}_2, s_1^2$ and s_2^2 are the sample means and variances of two random samples of sizes n_1 and n_2, respectively, from two independent normal populations with unknown but equal variances, then a **$100(1 - \alpha)\%$ confidence interval on the difference in means $\mu_1 - \mu_2$ is**

$$\bar{x}_1 - \bar{x}_2 - t_{\alpha/2, n_1+n_2-2}\, s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$\leq \mu_1 - \mu_2 \leq \bar{x}_1 - \bar{x}_2 + t_{\alpha/2, n_1+n_2-2}\, s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \qquad (10\text{-}19)$$

where $s_p = \sqrt{[(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2]/(n_1 + n_2 - 2)}$ is the pooled estimate of the common population standard deviation, and $t_{\alpha/2, n_1+n_2-2}$ is the upper $\alpha/2$ percentage point of the t distribution with $n_1 + n_2 - 2$ degrees of freedom.

10-2 Inference for a Difference in Means of Two Normal Distributions, Variances Unknown

Case 1: $\sigma_1^2 = \sigma_2^2 = \sigma^2$

Example 10-8

An article in the journal *Hazardous Waste and Hazardous Materials* (Vol. 6, 1989) reported the results of an analysis of the weight of calcium in standard cement and cement doped with lead. Reduced levels of calcium would indicate that the hydration mechanism in the cement is blocked and would allow water to attack various locations in the cement structure. Ten samples of standard cement had an average weight percent calcium of $\bar{x}_1 = 90.0$, with a sample standard deviation of $s_1 = 5.0$, while 15 samples of the lead-doped cement had an average weight percent calcium of $\bar{x}_2 = 87.0$, with a sample standard deviation of $s_2 = 4.0$.

We will assume that weight percent calcium is normally distributed and find a 95% confidence interval on the difference in means, $\mu_1 - \mu_2$, for the two types of cement. Furthermore, we will assume that both normal populations have the same standard deviation.

10-2 Inference for a Difference in Means of Two Normal Distributions, Variances Unknown

Case 1: $\sigma_1^2 = \sigma_2^2 = \sigma^2$

Example 10-8 (Continued)

The pooled estimate of the common standard deviation is found using Equation 10-12 as follows:

$$
\begin{aligned}
s_p^2 &= \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \\
&= \frac{9(5.0)^2 + 14(4.0)^2}{10 + 15 - 2} \\
&= 19.52
\end{aligned}
$$

10-2 Inference for a Difference in Means of Two Normal Distributions, Variances Unknown

Case 1: $\sigma_1^2 = \sigma_2^2 = \sigma^2$

Example 10-8 (Continued)

Therefore, the pooled standard deviation estimate is $s_p = \sqrt{19.52} = 4.4$. The 95% confidence interval is found using Equation 10-19:

$$\bar{x}_1 - \bar{x}_2 - t_{0.025,23}\, s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \le \mu_1 - \mu_2 \le \bar{x}_1 - \bar{x}_2 + t_{0.025,23}\, s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

or upon substituting the sample values and using $t_{0.025,23} = 2.069$,

$$90.0 - 87.0 - 2.069(4.4)\sqrt{\frac{1}{10} + \frac{1}{15}} \le \mu_1 - \mu_2$$

$$\le 90.0 - 87.0 + 2.069(4.4)\sqrt{\frac{1}{10} + \frac{1}{15}}$$

10-2 Inference for a Difference in Means of Two Normal Distributions, Variances Unknown

Case 1: $\sigma_1^2 = \sigma_2^2 = \sigma^2$

Example 10-8 (Continued)

which reduces to

$$-0.72 \leq \mu_1 - \mu_2 \leq 6.72$$

Notice that the 95% confidence interval includes zero; therefore, at this level of confidence we cannot conclude that there is a difference in the means. Put another way, there is no evidence that doping the cement with lead affected the mean weight percent of calcium; therefore, we cannot claim that the presence of lead affects this aspect of the hydration mechanism at the 95% level of confidence.

10-2 Inference for a Difference in Means of Two Normal Distributions, Variances Unknown

10-2.3 Confidence Interval on the Difference in Means, Variance Unknown

Case 2: $\sigma_1^2 \neq \sigma_2^2$

If $\bar{x}_1$, $\bar{x}_2$, s_1^2, and s_2^2 are the means and variances of two random samples of sizes n_1 and n_2, respectively, from two independent normal populations with unknown and unequal variances, an approximate $100(1 - \alpha)\%$ confidence interval on the difference in means $\mu_1 - \mu_2$ is

$$\bar{x}_1 - \bar{x}_2 - t_{\alpha/2,\nu}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \leq \mu_1 - \mu_2 \leq \bar{x}_1 - \bar{x}_2 + t_{\alpha/2,\nu}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \quad (10\text{-}20)$$

where ν is given by Equation 10-16 and $t_{\alpha/2,\nu}$ is the upper $\alpha/2$ percentage point of the t distribution with ν degrees of freedom.

10-4 Paired *t*-Test

- A special case of the two-sample t-tests of Section 10-2 occurs when the observations on the two populations of interest are collected in **pairs**.

- Each pair of observations, say (X_{1j}, X_{2j}), is taken under homogeneous conditions, but these conditions may change from one pair to another.

- The test procedure consists of analyzing the **differences** on each pair.

10-4 Paired *t*-Test

The Paired *t*-Test

Null hypothesis: $H_0\colon \mu_D = \Delta_0$

Test statistic: $$T_0 = \frac{\overline{D} - \Delta_0}{S_D/\sqrt{n}} \tag{10-22}$$

$D_j = X_{1j} - X_{2j}$

$\bar{D}$ = sample average of D_j

S_D = sample standard dev. of D_j

Alternative Hypothesis	Rejection Region
$H_1\colon \mu_D \neq \Delta_0$	$t_0 > t_{\alpha/2,n-1}$ or $t_0 < -t_{\alpha/2,n-1}$
$H_1\colon \mu_D > \Delta_0$	$t_0 > t_{\alpha,n-1}$
$H_1\colon \mu_D < \Delta_0$	$t_0 < -t_{\alpha,n-1}$

10-4 Paired *t*-Test

Example 10-9

An article in the *Journal of Strain Analysis* (1983, Vol. 18, No. 2) compares several methods for predicting the shear strength for steel plate girders. Data for two of these methods, the Karlsruhe and Lehigh procedures, when applied to nine specific girders, are shown in Table 10-2. We wish to determine whether there is any difference (on the average) between the two methods.

Table 10-2 Strength Predictions for Nine Steel Plate Girders (Predicted Load/Observed Load)

Girder	Karlsruhe Method	Lehigh Method	Difference d_j
S1/1	1.186	1.061	0.119
S2/1	1.151	0.992	0.159
S3/1	1.322	1.063	0.259
S4/1	1.339	1.062	0.277
S5/1	1.200	1.065	0.138
S2/1	1.402	1.178	0.224
S2/2	1.365	1.037	0.328
S2/3	1.537	1.086	0.451
S2/4	1.559	1.052	0.507

10-4 Paired t-Test

Example 10-9

The eight-step procedure is applied as follows:

1. The parameter of interest is the difference in mean shear strength between the two methods, say, $\mu_D = \mu_1 - \mu_2 = 0$.
2. H_0: $\mu_D = 0$
3. H_1: $\mu_D \neq 0$
4. $\alpha = 0.05$
5. The test statistic is

$$t_0 = \frac{\bar{d}}{s_D/\sqrt{n}}$$

6. Reject H_0 if $t_0 > t_{0.025,8} = 2.306$ or if $t_0 < -t_{0.025,8} = -2.306$.

10-4 Paired *t*-Test

Example 10-9

7. Computations: The sample average and standard deviation of the differences d_j are $\bar{d} = 0.2736$ and $s_D = 0.1356$, so the test statistic is

$$\boxed{t_0 = \frac{\bar{d}}{s_D/\sqrt{n}} = \frac{0.2736}{0.1356/\sqrt{9}} = 6.05}$$

8. Conclusions: Since $t_0 = 6.05 > 2.306$, we conclude that the strength prediction methods yield different results. Specifically, the data indicate that the Karlsruhe method produces, on the average, higher strength predictions than does the Lehigh method. The *P*-value for $t_0 = 6.05$ is $P = 0.0002$, so the test statistic is well into the critical region.

10-4 Paired *t*-Test

Paired Versus Unpaired Comparisons

So how do we decide to conduct the experiment? Should we pair the observations or not? Although there is no general answer to this question, we can give some guidelines based on the above discussion.

1. If the experimental units are relatively homogeneous (small σ) and the correlation within pairs is small, the gain in precision attributable to pairing will be offset by the loss of degrees of freedom, so an independent-sample experiment should be used.
2. If the experimental units are relatively heterogeneous (large σ) and there is large positive correlation within pairs, the paired experiment should be used. Typically, this case occurs when the experimental units are the *same* for both treatments; as in Example 10-9, the same girders were used to test the two methods.

10-4 Paired t-Test

A Confidence Interval for μ_D

Definition

If $\bar{d}$ and s_D are the sample mean and standard deviation of the difference of n random pairs of normally distributed measurements, a **$100(1-\alpha)\%$ confidence interval on the difference in means $\mu_D = \mu_1 - \mu_2$** is

$$\bar{d} - t_{\alpha/2,n-1} s_D/\sqrt{n} \le \mu_D \le \bar{d} + t_{\alpha/2,n-1} s_D/\sqrt{n} \qquad (10\text{-}23)$$

where $t_{\alpha/2,n-1}$ is the upper $\alpha/2\%$ point of the t-distribution with $n-1$ degrees of freedom.

10-4 Paired t-Test

Example 10-10

Table 10-3 Time in Seconds to Parallel Park Two Automobiles

Subject	Automobile 1(x_{1j})	Automobile 2(x_{2j})	Difference (d_j)
1	37.0	17.8	19.2
2	25.8	20.2	5.6
3	16.2	16.8	−0.6
4	24.2	41.4	−17.2
5	22.0	21.4	0.6
6	33.4	38.4	−5.0
7	23.8	16.8	7.0
8	58.2	32.2	26.0
9	33.6	27.8	5.8
10	24.4	23.2	1.2
11	23.4	29.6	−6.2
12	21.2	20.6	0.6
13	36.2	32.2	4.0
14	29.8	53.8	−24.0

10-4 Paired t-Test

Example 10-10

The journal *Human Factors* (1962, pp. 375-380) reports a study in which $n = 14$ subjects were asked to parallel park two cars having very different wheel bases and turning radii. The time in seconds for each subject was recorded and is given in Table 10-3. From the column of observed differences we calculate $\bar{d} = 1.21$ and $s_D = 12.68$. The 90% confidence interval for $\mu_D = \mu_1 - \mu_2$ is found from Equation 9-24 as follows:

$$\bar{d} - t_{0.05,13}s_D/\sqrt{n} \le \mu_D \le \bar{d} + t_{0.05,13}s_D/\sqrt{n}$$

$$1.21 - 1.771(12.68)/\sqrt{14} \le \mu_D \le 1.21 + 1.771(12.68)/\sqrt{14}$$

$$-4.79 \le \mu_D \le 7.21$$

Notice that the confidence interval on μ_D includes zero. This implies that, at the 90% level of confidence, the data do not support the claim that the two cars have different mean parking times μ_1 and μ_2. That is, the value $\mu_D = \mu_1 - \mu_2 = 0$ is not inconsistent with the observed data.

10-5 Inferences on the Variances of Two Normal Populations

10-5.1 The *F* Distribution

We wish to test the hypotheses:

$$H_0: \sigma_1^2 = \sigma_2^2$$
$$H_1: \sigma_1^2 \neq \sigma_2^2$$

- The development of a test procedure for these hypotheses requires a new probability distribution, the *F* distribution.

10-5 Inferences on the Variances of Two Normal Populations

10-5.1 The *F* Distribution

Let W and Y be independent chi-square random variables with u and v degrees of freedom, respectively. Then the ratio

$$\boxed{F = \frac{W/u}{Y/v}} \tag{10-26}$$

has the probability density function

$$f(x) = \frac{\Gamma\left(\frac{u+v}{2}\right)\left(\frac{u}{v}\right)^{u/2} x^{(u/2)-1}}{\Gamma\left(\frac{u}{2}\right)\Gamma\left(\frac{v}{2}\right)\left[\left(\frac{u}{v}\right)x + 1\right]^{(u+v)/2}}, \qquad 0 < x < \infty \tag{10-27}$$

and is said to follow the F distribution with u degrees of freedom in the numerator and v degrees of freedom in the denominator. It is usually abbreviated as $F_{u,v}$.

10-5 Inferences on the Variances of Two Normal Populations

10-5.1 The *F* Distribution

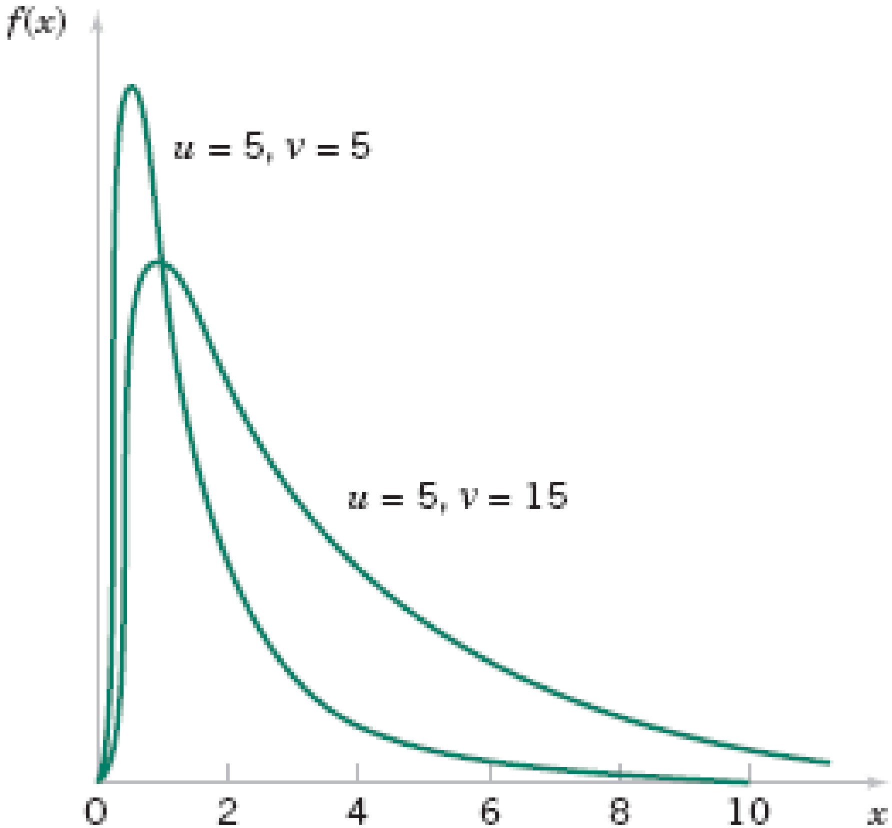


Figure 10-4 Probability density functions of two *F* distributions.

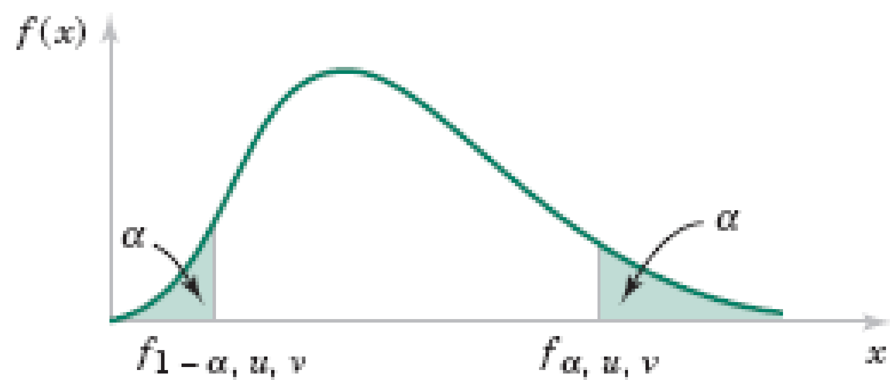


Figure 10-5 Upper and lower percentage points of the *F* distribution.

10-5 Inferences on the Variances of Two Normal Populations

10-5.1 The *F* Distribution

The lower-tail percentage points $f_{1-\alpha,u,\nu}$ can be found as follows.

$$f_{1-\alpha,u,v} = \frac{1}{f_{\alpha,v,u}} \tag{10-28}$$

10-5 Inferences on the Variances of Two Normal Populations

10-5.2 Hypothesis Tests on the Ratio of Two Variances

Let $X_{11}, X_{12}, \ldots, X_{1n_1}$ be a random sample from a normal population with mean μ_1 and variance σ_1^2, and let $X_{21}, X_{22}, \ldots, X_{2n_2}$ be a random sample from a second normal population with mean μ_2 and variance σ_2^2. Assume that both normal populations are independent. Let S_1^2 and S_2^2 be the sample variances. Then the ratio

$$F = \frac{(n_1-1)S_1^2/\sigma_1^2(n_1-1)}{(n_2-1)S_2^2/\sigma_2^2(n_2-1)} \qquad \boxed{F = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}}$$

has an F distribution with $n_1 - 1$ numerator degrees of freedom and $n_2 - 1$ denominator degrees of freedom.

10-5 Inferences on the Variances of Two Normal Populations

10-5.2 Hypothesis Tests on the Ratio of Two Variances

Null hypothesis: $\boxed{H_0\colon \sigma_1^2 = \sigma_2^2}$

Test statistic: $$F_0 = \frac{S_1^2}{S_2^2} \tag{10-29}$$

Alternative Hypotheses	Rejection Criterion
$H_1\colon \sigma_1^2 \neq \sigma_2^2$	$f_0 > f_{\alpha/2,n_1-1,n_2-1}$ or $f_0 < f_{1-\alpha/2,n_1-1,n_2-1}$
$H_1\colon \sigma_1^2 > \sigma_2^2$	$f_0 > f_{\alpha,n_1-1,n_2-1}$
$H_1\colon \sigma_1^2 < \sigma_2^2$	$f_0 < f_{1-\alpha,n_1-1,n_2-1}$

10-5 Inferences on the Variances of Two Normal Populations

Example 10-11

Oxide layers on semiconductor wafers are etched in a mixture of gases to achieve the proper thickness. The variability in the thickness of these oxide layers is a critical characteristic of the wafer, and low variability is desirable for subsequent processing steps. Two different mixtures of gases are being studied to determine whether one is superior in reducing the variability of the oxide thickness. Twenty wafers are etched in each gas. The sample standard deviations of oxide thickness are $s_1 = 1.96$ angstroms and $s_2 = 2.13$ angstroms, respectively. Is there any evidence to indicate that either gas is preferable? Use $\alpha = 0.05$.

10-5 Inferences on the Variances of Two Normal Populations

Example 10-11

The eight-step hypothesis-testing procedure may be applied to this problem as follows:

1. The parameters of interest are the variances of oxide thickness σ_1^2 and σ_2^2. We will assume that oxide thickness is a normal random variable for both gas mixtures.
2. H_0: $\sigma_1^2 = \sigma_2^2$
3. H_1: $\sigma_1^2 \neq \sigma_2^2$
4. $\alpha = 0.05$
5. The test statistic is given by Equation 10-29:

$$f_0 = \frac{s_1^2}{s_2^2}$$

10-5 Inferences on the Variances of Two Normal Populations

Example 10-11

6. $\boxed{\text{Since } n_1 = n_2 = 20, \text{ we will reject } H_0\text{: } \sigma_1^2 = \sigma_2^2 \text{ if } f_0 > f_{0.025,19,19} = 2.53 \text{ or if } f_0 < f_{0.975,19,19} = 1/f_{0.025,19,19} = 1/2.53 = 0.40.}$

7. Computations: Since $s_1^2 = (1.96)^2 = 3.84$ and $s_2^2 = (2.13)^2 = 4.54$, the test statistic is

$$\boxed{f_0 = \frac{s_1^2}{s_2^2} = \frac{3.84}{4.54} = 0.85}$$

8. Conclusions: Since $f_{0.975,19,19} = 0.40 < f_0 = 0.85 < f_{0.025,19,19} = 2.53$, we <u>cannot reject the null hypothesis</u> H_0: $\sigma_1^2 = \sigma_2^2$ at the 0.05 level of significance. Therefore, there is no strong evidence to indicate that either gas results in a smaller variance of oxide thickness.

10-5 Inferences on the Variances of Two Normal Populations

10-5.3 Type II Error and Choice of Sample Size

Appendix Charts VII*o*, VII*p*, VII*q*, and VII*r* provide operating characteristic curves for the *F*-test given in Section 10-5.1 for $\alpha = 0.05$ and $\alpha = 0.01$, assuming that $n_1 = n_2 = n$. Charts VII*o* and VII*p* are used with the two-sided alternate hypothesis. They plot β against the abscissa parameter

$$\boxed{\lambda = \frac{\sigma_1}{\sigma_2}} \qquad (10\text{-}30)$$

for various $n_1 = n_2 = n$. Charts VII*q* and VII*r* are used for the one-sided alternative hypotheses.

10-5 Inferences on the Variances of Two Normal Populations

Example 10-12

For the semiconductor wafer oxide etching problem in Example 10-11, suppose that one gas resulted in a standard deviation of oxide thickness that is half the standard deviation of oxide thickness of the other gas. If we wish to detect such a situation with probability at least 0.80, is the sample size $n_1 = n_2 = 20$ adequate?

Note that if one standard deviation is half the other,

$$\lambda = \frac{\sigma_1}{\sigma_2} = 2$$

By referring to Appendix Chart VII*o* with $n_1 = n_2 = n = 20$ and $\lambda = 2$, we find that $\beta \simeq 0.20$. Therefore, if $\beta = 0.20$, the power of the test (which is the probability that the difference in standard deviations will be detected by the test) is 0.80, and we conclude that the sample sizes $n_1 = n_2 = 20$ are adequate.

10-5 Inferences on the Variances of Two Normal Populations

10-5.4 Confidence Interval on the Ratio of Two Variances

If s_1^2 and s_2^2 are the sample variances of random samples of sizes n_1 and n_2, respectively, from two independent normal populations with unknown variances σ_1^2 and σ_2^2, then a **$100(1-\alpha)\%$ confidence interval on the ratio σ_1^2/σ_2^2** is

$$\boxed{\frac{s_1^2}{s_2^2} f_{1-\alpha/2,\underline{n_2-1,n_1-1}} \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \frac{s_1^2}{s_2^2} f_{\alpha/2,\underline{n_2-1,n_1-1}}} \qquad (10\text{-}31)$$

where $f_{\alpha/2,n_2-1,n_1-1}$ and $f_{1-\alpha/2,n_2-1,n_1-1}$ are the upper and lower $\alpha/2$ percentage points of the F distribution with $n_2 - 1$ numerator and $n_1 - 1$ denominator degrees of freedom, respectively. A confidence interval on the ratio of the standard deviations can be obtained by taking square roots in Equation 10-31.

10-5 Inferences on the Variances of Two Normal Populations

Example 10-13

A company manufactures impellers for use in jet-turbine engines. One of the operations involves grinding a particular surface finish on a titanium alloy component. Two different grinding processes can be used, and both processes can produce parts at identical mean surface roughness. The manufacturing engineer would like to select the process having the least variability in surface roughness. A random sample of $n_1 = 11$ parts from the first process results in a sample standard deviation $s_1 = 5.1$ microinches, and a random sample of $n_2 = 16$ parts from the second process results in a sample standard deviation of $s_2 = 4.7$ microinches. We will find a 90% confidence interval on the ratio of the two standard deviations, σ_1/σ_2.

10-5 Inferences on the Variances of Two Normal Populations

Example 10-13

Assuming that the two processes are independent and that surface roughness is normally distributed, we can use Equation 10-31 as follows:

$n_1=11$
$n_2=16$

$$\frac{s_1^2}{s_2^2} f_{0.95,15,10} \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{s_1^2}{s_2^2} f_{0.05,15,10}$$

$$\frac{(5.1)^2}{(4.7)^2} 0.39 \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{(5.1)^2}{(4.7)^2} 2.85$$

$$0.678 \le \frac{\sigma_1}{\sigma_2} \le 1.832$$

[it includes 1, so cannot claim that the two standard deviations are different at 90% confidence.]

where $f_{0.95,15,10} = {}^{1}\!\big/_{f_{0.05,10,15}} = {}^{1}\!\big/_{2.54} = 0.39$

10-6 Inference on Two Population Proportions

10-6.1 Large-Sample Test on the Difference in Population Proportions

We wish to test the hypotheses:

$$H_0: p_1 = p_2$$
$$H_1: p_1 \neq p_2$$

10-6 Inference on Two Population Proportions

10-6.1 Large-Sample Test on the Difference in Population Proportions

The following test statistic is distributed approximately as standard normal and is the basis of the test:

$$Z = \frac{\hat{P}_1 - \hat{P}_2 - (p_1 - p_2)}{\sqrt{\frac{p_1(1 - p_1)}{n_1} + \frac{p_2(1 - p_2)}{n_2}}} \tag{10-32}$$

10-6 Inference on Two Population Proportions

10-6.1 Large-Sample Test on the Difference in Population Proportions

Null hypothesis: $H_0\colon p_1 = p_2$ where $\hat{P} = \dfrac{X_1 + X_2}{n_1 + n_2}$

Test statistic:

$$Z_0 = \frac{\hat{P}_1 - \hat{P}_2}{\sqrt{\hat{P}(1 - \hat{P})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \qquad (10\text{-}33)$$

Alternative Hypotheses	Rejection Criterion
$H_1\colon p_1 \neq p_2$	$z_0 > z_{\alpha/2}$ or $z_0 < -z_{\alpha/2}$
$H_1\colon p_1 > p_2$	$z_0 > z_\alpha$
$H_1\colon p_1 < p_2$	$z_0 < -z_\alpha$

10-6 Inference on Two Population Proportions

Example 10-14

Extracts of St. John's Wort are widely used to treat depression. An article in the April 18, 2001 issue of the *Journal of the American Medical Association* ("Effectiveness of St. John's Wort on Major Depression: A Randomized Controlled Trial") compared the efficacy of a standard extract of St. John's Wort with a placebo in 200 outpatients diagnosed with major depression. Patients were randomly assigned to two groups; one group received the St. John's Wort, and the other received the placebo. After eight weeks, 19 of the placebo-treated patients showed improvement, whereas 27 of those treated with St. John's Wort improved. Is there any reason to believe that St. John's Wort is effective in treating major depression? Use $\alpha = 0.05$.
The eight-step hypothesis testing procedure leads to the following results:

Placebo: yalancı ilaç
St. John's wort: sarı kantaron

10-6 Inference on Two Population Proportions

Example 10-14

1. The parameters of interest are p_1 and p_2, the proportion of patients who improve following treatment with St. John's Wort (p_1) or the placebo (p_2).
2. H_0: $p_1 = p_2$
3. H_1: $p_1 \neq p_2$
4. $\alpha = 0.05$
5. The test statistic is

$$z_0 = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where $\hat{p}_1 = 27/100 = 0.27$, $\hat{p}_2 = 19/100 = 0.19$, $n_1 = n_2 = 100$, and

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{19 + 27}{100 + 100} = 0.23$$

10-6 Inference on Two Population Proportions

Example 10-14

6. $\boxed{\text{Reject } H_0\text{: } p_1 = p_2 \text{ if } z_0 > z_{0.025} = 1.96 \text{ or if } z_0 < -z_{0.025} = -1.96.}$

7. Computations: The value of the test statistic is

$$\boxed{z_0 = \frac{0.27 - 0.19}{\sqrt{0.23(0.77)\left(\frac{1}{100} + \frac{1}{100}\right)}} = 1.35}$$

8. Conclusions: Since $z_0 = 1.35$ does not exceed $z_{0.025}$, we cannot reject the null hypothesis. Note that the P-value is $P \simeq 0.177$. There is insufficient evidence to support the claim that St. John's Wort is effective in treating major depression.

10-6 Inference on Two Population Proportions

Minitab Output for Example 10-14

Minitab Computations

Test and CI for Two Proportions

Sample	X	N	Sample p
1	27	100	0.270000
2	19	100	0.190000

Estimate for p(1) − p(2): 0.08

95% CI for p(1) − p(2): (−0.0361186, 0.196119)

Test for p(1) − p(2) = 0 (vs not = 0): Z = 1.35 P-Value = 0.177

10-6 Inference on Two Population Proportions

10-6.2 Type II Error and Choice of Sample Size

If the alternative hypothesis is two sided, the β-error is

$$\beta = \Phi\left[\frac{z_{\alpha/2}\sqrt{\bar{p}\bar{q}(1/n_1 + 1/n_2)} - (p_1 - p_2)}{\sigma_{\hat{P}_1 - \hat{P}_2}}\right] - \Phi\left[\frac{-z_{\alpha/2}\sqrt{\bar{p}\bar{q}(1/n_1 + 1/n_2)} - (p_1 - p_2)}{\sigma_{\hat{P}_1 - \hat{P}_2}}\right] \quad (10\text{-}35)$$

$$\sigma_{\hat{P}_1 - \hat{P}_2} = \sqrt{\frac{p_1(1 - p_1)}{n_1} + \frac{p_2(1 - p_2)}{n_2}}$$

$$\bar{p} = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2} \quad \text{and} \quad \bar{q} = \frac{n_1(1 - p_1) + n_2(1 - p_2)}{n_1 + n_2}$$

10-6 Inference on Two Population Proportions

10-6.2 Type II Error and Choice of Sample Size

If the alternative hypothesis is $H_1\colon p_1 > p_2$,

$$\beta = \Phi\left[\frac{z_\alpha\sqrt{\bar{p}\bar{q}(1/n_1 + 1/n_2)} - (p_1 - p_2)}{\sigma_{\hat{P}_1-\hat{P}_2}}\right] \tag{10-36}$$

and if the alternative hypothesis is $H_1\colon p_1 < p_2$,

$$\beta = 1 - \Phi\left[\frac{-z_\alpha\sqrt{\bar{p}\bar{q}(1/n_1 + 1/n_2)} - (p_1 - p_2)}{\sigma_{\hat{P}_1-\hat{P}_2}}\right] \tag{10-37}$$

10-6 Inference on Two Population Proportions

10-6.2 Type II Error and Choice of Sample Size

For the two-sided alternative, the common sample size is $n_1 = n_2 = n$

$$n = \frac{\left[z_{\alpha/2}\sqrt{(p_1 + p_2)(q_1 + q_2)/2} + z_\beta\sqrt{p_1 q_1 + p_2 q_2}\right]^2}{(p_1 - p_2)^2} \tag{10-38}$$

where $q_1 = 1 - p_1$ and $q_2 = 1 - p_2$.

10-6 Inference on Two Population Proportions

10-6.3 Confidence Interval on the Difference in the Population Proportions

If $\hat{p}_1$ and $\hat{p}_2$ are the sample proportions of observation in two independent random samples of sizes n_1 and n_2 that belong to a class of interest, **an approximate two-sided $100(1-\alpha)\%$ confidence interval on the difference in the true proportions $p_1 - p_2$ is**

$$\hat{p}_1 - \hat{p}_2 - z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

$$\leq p_1 - p_2 \leq \hat{p}_1 - \hat{p}_2 + z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} \qquad (10\text{-}39)$$

where $z_{\alpha/2}$ is the upper $\alpha/2$ percentage point of the standard normal distribution.

10-6 Inference on Two Population Proportions

Example 10-15

Consider the process manufacturing crankshaft bearings described in Example 8-6. Suppose that a modification is made in the surface finishing process and that, subsequently, a second random sample of 85 axle shafts is obtained. The number of defective shafts in this second sample is 8. Therefore, since $n_1 = 85$, $\hat{p}_1 = 0.12$, $n_2 = 85$, and $\hat{p}_2 = 8/85 = 0.09$, we can obtain an approximate 95% confidence interval on the difference in the proportion of defective bearings produced under the two processes from Equation 10-39 as follows:

$$\hat{p}_1 - \hat{p}_2 - z_{0.025}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

$$\leq p_1 - p_2 \leq \hat{p}_1 - \hat{p}_2 + z_{0.025}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

10-6 Inference on Two Population Proportions

Example 10-15

or

$$0.12 - 0.09 - 1.96\sqrt{\frac{0.12(0.88)}{85} + \frac{0.09(0.91)}{85}} \le p_1 - p_2 \le 0.12 - 0.09 + 1.96\sqrt{\frac{0.12(0.88)}{85} + \frac{0.09(0.91)}{85}}$$

This simplifies to

$$-0.06 \le p_1 - p_2 \le 0.12$$

This confidence interval includes zero, so, based on the sample data, it seems unlikely that the changes made in the surface finish process have reduced the proportion of defective crankshaft bearings being produced.