Design of Experiments
Part 2– Basic Statistical Concepts (2)

- Simple **comparative** experiments
  - Review the independent two-sample t-test
  - Comparing a single mean to a specified value
  - The paired two-sample t-test
  - Inferences about the variances of normal distribution

Average breaking strength of fabric (page 46)

A supplier submits lots of fabric to a textile manufacturer. The customer wants to know if the lot average breaking strength exceeds 200 psi. If so, she wants to accept the lot. Past experience indicates that a reasonable value for the variance of breaking strength is 100 $(psi)^2$. Four lots are randomly selected, and the average breaking strength observed is

$\bar{y}$ =214psi. Should she accept the lots?

The Hypothesis Testing Framework (when variance is known)

- Assumption: Sampling from a **normal** distribution
- Statistical hypotheses: $H_0: \mu = 200$

  $H_1: \mu > 200$
- Test Statistic:

$$Z_0 = \frac{\bar{y} - u_0}{\sigma / \sqrt{n}} \overset{\text{Under } H_0}{\sim} N(0,1)$$

$$= \frac{214 - 200}{10 / \sqrt{4}} = 2.8$$

- Reject region or P-value: for $\alpha=0.05$, $Z_\alpha = Z_{0.05} = 1.645$ (Appendix Table I). Reject if $Z_0 > Z_{0.05}$ or if p-value<0.05.
- Conclusion: Since $Z_0 > 1.645$ or p-value= 0.002555130 (from R output),

H_0 is rejected, and we conclude that the lot average breaking strength exceeds 200psi. Therefore she can accept the lot

Obtaining critical value or p-value

- Critical value: Appendix Table I
- P-value : R output
    - R code:

```
> 1-pnorm(2.8)
[1] 0.002555130
```

Confidence Interval

(when variance is known

and H_a is two-sided)

$$\bar{y} - z_{\alpha/2}\sigma/\sqrt{n} \le \mu \le \bar{y} + z_{\alpha/2}\sigma/\sqrt{n}$$

$z_{\alpha/2}$ can be obtained from

- Appendix Table I
- R output

R code: >qnorm(1-α/2)

Q: How to obtain CI when variance is known and Ha is one-sided?

The Hypothesis Testing Framework (when variance is unknown)

- Assumption: Sampling from a **normal** distribution
- Statistical hypotheses: $H_0 : \mu = \mu_0$

  $H_1 : \mu \neq (>,<)\ \mu_0$
- Test Statistic:

$$t_0 = \frac{\bar{y} - u_0}{s / \sqrt{n}} \overset{\text{Under } H_0}{\sim} t(n-1)$$

- Reject region or P-value: for α=0.05, critical value t can be found from Appendix Table II. Reject H_0 if $|t_0|>t$ ($t_0>t$, $t_0<t$) or if p-value<0.05
- Conclusion

Obtaining critical value or p-value

- **Critical value**: Appendix Table II
- **P-value** : R output
  - R code:

    >1-pt(t_0 ,n-1) # if one-sided alternative (>)

    > pt(t_0 ,n-1) # if one-sided alternative (<)

    > 2*(1-pt(t_0 ,n-1)) # if two-sided alternative and t_0 >0

    > 2*pt(t_0 ,n-1) # if two-sided alternative and t_0 <0

Confidence Interval

(when variance is unknown

and H_a is two-sided)

$$\bar{y} - t_{\alpha/2,n-1} s/\sqrt{n} \le \mu \le \bar{y} + t_{\alpha/2,n-1} s/\sqrt{n}$$

$t_{\alpha/2,n-1}$ can be obtained from

- Appendix Table II
- R output

R code: >qt(1-α/2,n-1)

Q: How to obtain CI when variance is unknown and Ha is one-sided?

Hardness Testing Experiment (page 49)

Specimen	Tip 1	Tip2
1	7	6
2	3	3
3	3	5
4	4	3
5	8	8
6	3	3
7	2	4
8	9	9
9	5	4
10	4	5

- Press a rod with two different tips into a metal specimen to test the hardness of the specimen.
- The hardness is determined by the depth of the depression
- Specimens are not exactly homogeneous in some other way that might affect the hardness

*Q: **Does one tip produce different hardness readings than the other? Is completely randomized design and associated two-sample t-test suitable?***

Hardness Testing Experiment (page 49)

Specimen	Tip 1	Tip2	diff (d_j)
1	7	6	1
2	3	3	0
3	3	5	-2
4	4	3	1
5	8	8	0
6	3	3	0
7	2	4	-2
8	9	9	0
9	5	4	1
10	4	5	-1
Mean(Y)	4.8	4.9	-0.1
SD(Y)	2.39	2.2	1.197

By pairing, the estimate of variability is reduced by nearly 50%

A completely randomized design for this problem will

- increase the variability of the hardness measurements
- tend to inflate the experimental error
- make a true difference harder to detect

It is beneficial if two hardness determination can be made on the same specimen

The Hypothesis Testing Framework (when variance is unknown)

- Assumption: d_js are sampled from a **normal** distribution $N(\mu_d , \sigma_d^2)$
- Statistical hypotheses: $H_0 : \mu_d = 0$

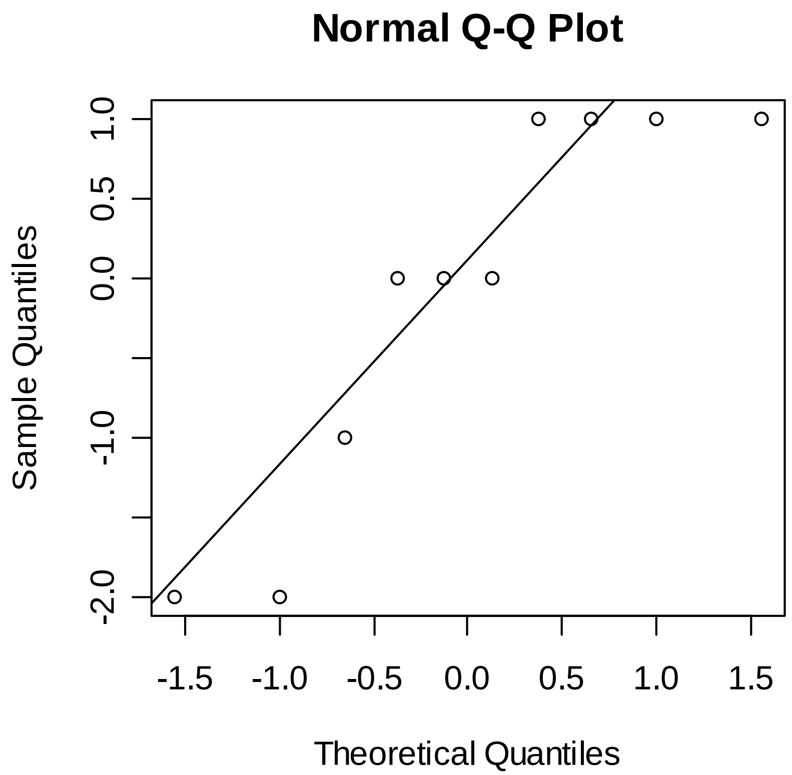
  $H_1 : \mu_d \neq 0$
- Test Statistic:

$$t_0 = \frac{\bar{d}}{s_d / \sqrt{n}} \overset{\text{Under } H_0}{\sim} t(n-1)$$

$$= \frac{-0.1}{1.197 / \sqrt{10}} = -0.26$$

- Reject region or P-value: for α=0.05, critical value t =2.262 (Appendix Table II (or R output)). Reject H_0 if $|t_0|>2.262$ or p-value <0.05
- Conclusion: because $|t_0|=0.26<2.262$ or p-value=0.80 (from R output), we cannot reject the hypothesis H_0 . Therefore, there is no evidence to indicate that the two tips produce different hardness readings.

Checking Assumptions – The Normal Probability Plot



R Paired *t*-Test Results

```
> t.test(tip1, tip2, paired=T)
 Paired t-test

data:  tip1 and tip2
t = -0.2641, df = 9, p-value = 0.7976
alternative hypothesis: true difference
in means is not equal to 0
95 percent confidence interval:
 -0.9564389  0.7564389
sample estimates:
mean of the differences
                   -0.1
```

Warning:

Blocking (or pairing when considering two samples) is not always the best design strategy. If the within-block variability is the same as the between-block variability, blocking would be a poor choice of design which results in the loss of n-1 degree of freedom and lead to a wider confidence interval on the difference between the two samples.

The Hypothesis Testing Framework on the variance of normal distribution

- Assumption: Sampling from a **normal** distribution
- Statistical hypotheses: $H_0: \quad \sigma^2 = \sigma_0^2$

  $H_1: \quad \sigma^2 \neq \sigma_0^2$
- Test Statistic:

$$\chi_0^2 = \frac{(n-1)s^2}{\sigma_0^2} \overset{\text{Under } H_0}{\sim} \chi^2(n-1)$$

- Reject region or P-value: Reject if $\chi_0^2 > \chi^2_{\alpha/2,n-1}$ or $\chi_0^2 < \chi^2_{1-\alpha/2,n-1}$
- Conclusion

Confidence Interval:

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2,n-1}} \leq \sigma^2 \leq \frac{(n-1)s^2}{\chi^2_{1-\alpha/2,n-1}}$$

Obtaining critical value or p-value

- Critical value: Appendix Table III
- P-value : R output

```
>1-pchisq(, n-1)
```

The Hypothesis Testing Framework on the equivalence of variances of two normal distribution

- Assumption: Both samples from **normal** distributions
- Statistical hypotheses: $H_0: \quad \sigma_1^2 = \sigma_2^2$

  $H_1: \quad \sigma_1^2 \neq \sigma_2^2$
- Test Statistic:

$$F_0 = \frac{s_1^2}{s_2^2} \overset{\text{Under } H_0}{\sim} F(n_1 - 1, n_2 - 1)$$

- Reject region or P-value: Reject if $F_0 > F_{\alpha/2, n_1-1, n_2-1}$ or $F_0 < F_{1-\alpha/2, n_1-1, n_2-1}$ or if p-value<α
- Conclusion

Confidence Interval:

$$\frac{s_1^2}{s_2^2} F_{1-\alpha/2, n_2-1, n_1-1} \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \frac{s_1^2}{s_2^2} F_{\alpha/2, n_2-1, n_1-1}$$

Obtaining critical value or p-value

- Critical value: Appendix Table IV
- P-value : R output

If $F_0>1$, use the following R code
$>1\text{-pf}(F_0, n_1-1, n_2-1)+\text{pf}(1/F_0, n_1-1, n_2-1)$

If $F_0<1$, use the following R code
$>1\text{-pf}(1/F_0, n_1-1, n_2-1)+\text{pf}(F_0, n_1-1, n_2-1)$

Example 2-2 (page 54)

- A chemical engineer wants to test if the old equipment, type 1, has a larger variance than the new one. Two random samples of $n_1=12$ and $n_2=10$ observations are taken, and the two sample variances are 14.5 and 10.8 respectively.
  - Statistical hypotheses: $H_0: \ \sigma_1^2 = \sigma_2^2$
    $H_1: \ \sigma_1^2 \neq \sigma_2^2$
  - Test Statistic: $$F_0 = \frac{s_1^2}{s_2^2} = \frac{14.5}{10.8} = 1.34$$
  - Reject region or P-value: critical value=$F_{0.05,11,9}$=3.10
    p-value=0.654
  - Conclusion: since $F_0 < 3.10$ or p-value >0.05, the null hypothesis cannot be rejected. We have found insufficient statistical evidence to conclude that the variance of the old equipment is greater than the variance of the new equipment.