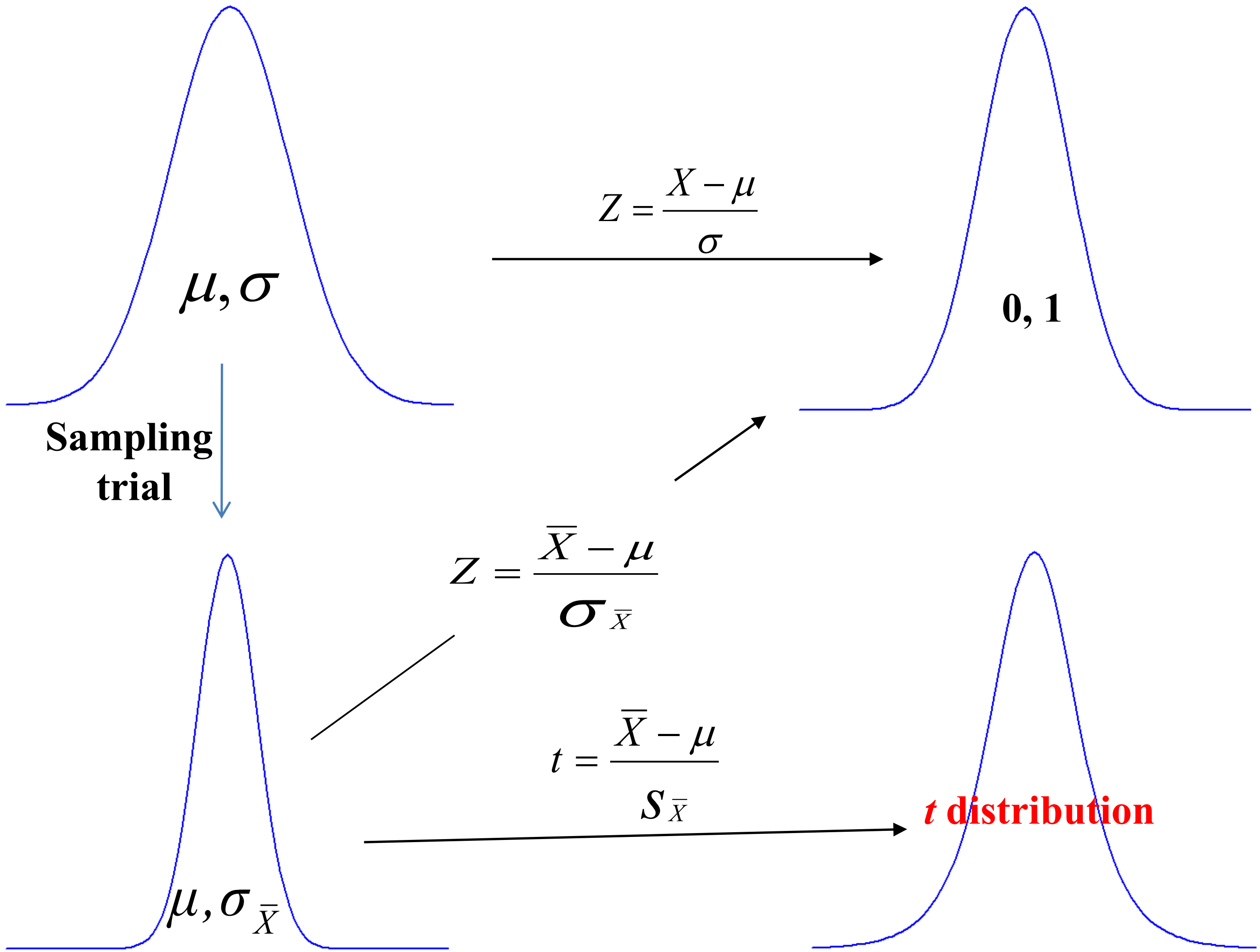
t test

Student *t* distribution



Characteristics of *t* distribution curve

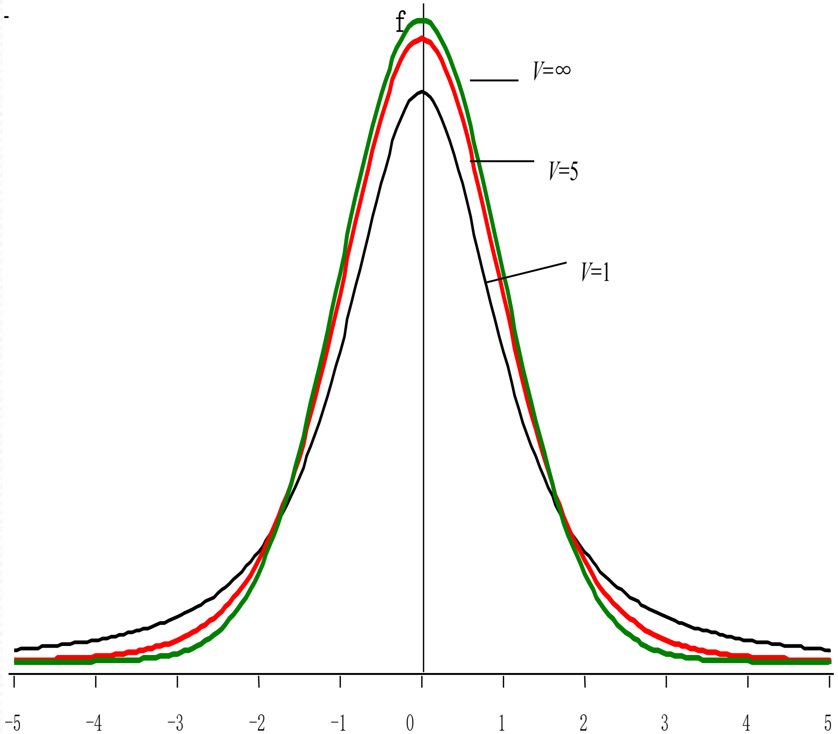


Figure 4-3 t-distribution graphs in different df

- It is symmetrical by y axis and has one apex;
- Only one parameter, ν(degree of freedom, ν=n-1) determines the shape of t-distribution.
- The total area under the t-distribution equals to 1. The area under the curve between any of interval (t1 and t2) can be found by calculus

Who introduced *t* test?

William Sealy Gosset

Student in 1908

t **test, also known as Student's** ***t*** **test, was invented by** ***W. S. Gossett [1876-1937] to handle small samples for quality control in brewing.***

What is t test used for?

A common hypothesis test for inferring whether two population means are equal, ***especially in the cases of small sample sizes.***

Contents

1. One sample *t* test
2. Paired *t* test
3. *t* test for two independent samples
4. Tests for normality and equal variances

1. One sample *t* test

1. One sample *t* test

One sample *t*-test is suitable for comparing the sample mean with the known population mean.

The testing condition includes the assumption of

- **Normal distribution, especially when n is small**

Example:

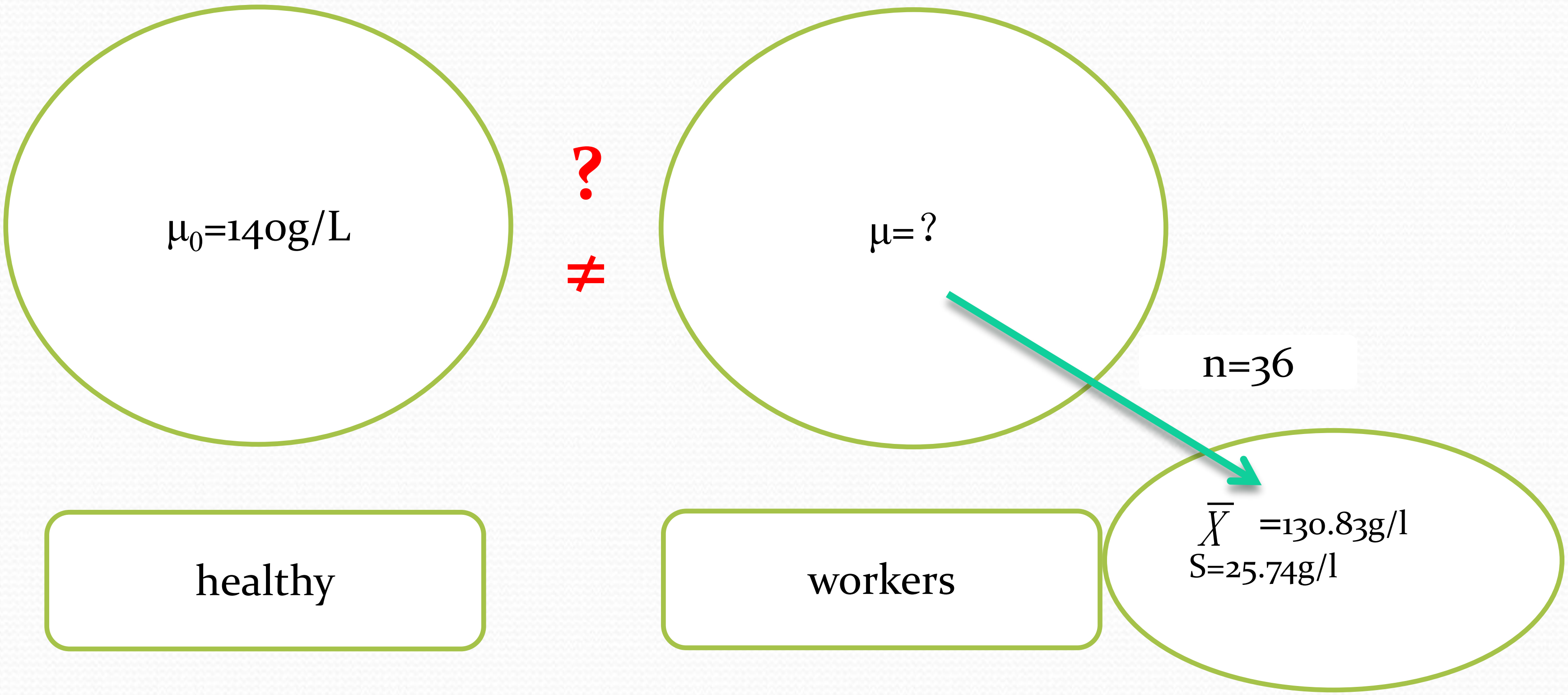
A doctor measured the hemoglobin concentrations in 36 workers engaged in lead operations, with

$$\overline{X} = 130.83 g / L, \quad S = 25.74 g / L$$

Is the mean of hemoglobin concentrations of lead workers smaller than that of the healthy individuals (μ_0=140g/L)?

assume the underlying distribution is normal

Question



We find that hemoglobin concentrations of lead workers obey normal distribution:

(1) Establish the hypothesis and determine the significance level

(a) H_0:

μ=140, the mean of hemoglobin concentrations of lead workers is equal to that of the healthy individuals (μ_0=140g/L).

(b) H_1:

μ<140, the mean of hemoglobin concentrations of lead workers is smaller than that of the healthy individuals (μ_0=140g/L).

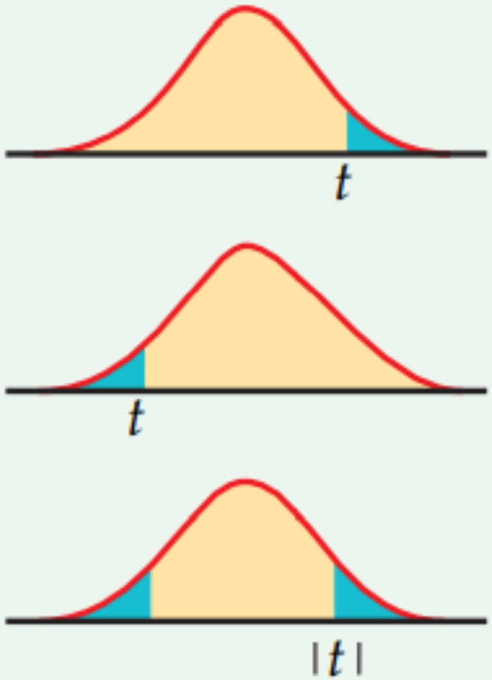
α =0.05

One tailed vs two tailed test

$H_a: \mu > \mu_0$ is $P(T \geq t)$

$H_a: \mu < \mu_0$ is $P(T \leq t)$

$H_a: \mu \neq \mu_0$ is $2P(T \geq |t|)$



One-Tailed Test	Two-Tailed Test
A test of any statistical hypothesis, where the alternative hypothesis is **one-tailed** either right-tailed or left-tailed.	A test of a statistical hypothesis, where the alternative hypothesis is **two-tailed**.
For one-tailed, we use either > or < sign for the alternative hypothesis.	For two-tailed, we use ≠ sign for the alternative hypothesis.
When the alternative hypothesis specifies a direction then we use a one-tailed test.	If no direction is given then we will use a two-tailed test.
Critical region lies entirely on either the right side or left side of the sampling distribution.	Critical region is given by the portion of the area lying in both the tails of the probability curve of the test statistic.

One-Tailed Test	Two-Tailed Test
Here, the Entire **level of significance** (α) i.e. 5% has either in the left tail or right tail.	It splits the **level of significance (α)** into half.
Rejection region is either from the left side or right side of the sampling distribution.	Rejection region is from both sides i.e. left and right of the sampling distribution.
It checks the relation between the variable in a singles direction.	It checks the relation between the variables in any direction.
It is used to check whether the one mean is different from another mean or not.	It is used to check whether the two mean different from one another or not.

two-tailed tests are much more common than one-tailed tests

The test statistic is

$$t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \qquad \nu = n - 1$$

Where S is the sample standard deviation, *n* is the sample size, *v* is the degree of freedom.

(2) Calculate the t test statistic

$$t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} = -2.138 \quad \nu = n\text{-}1 = 36\text{-}1 = 35$$

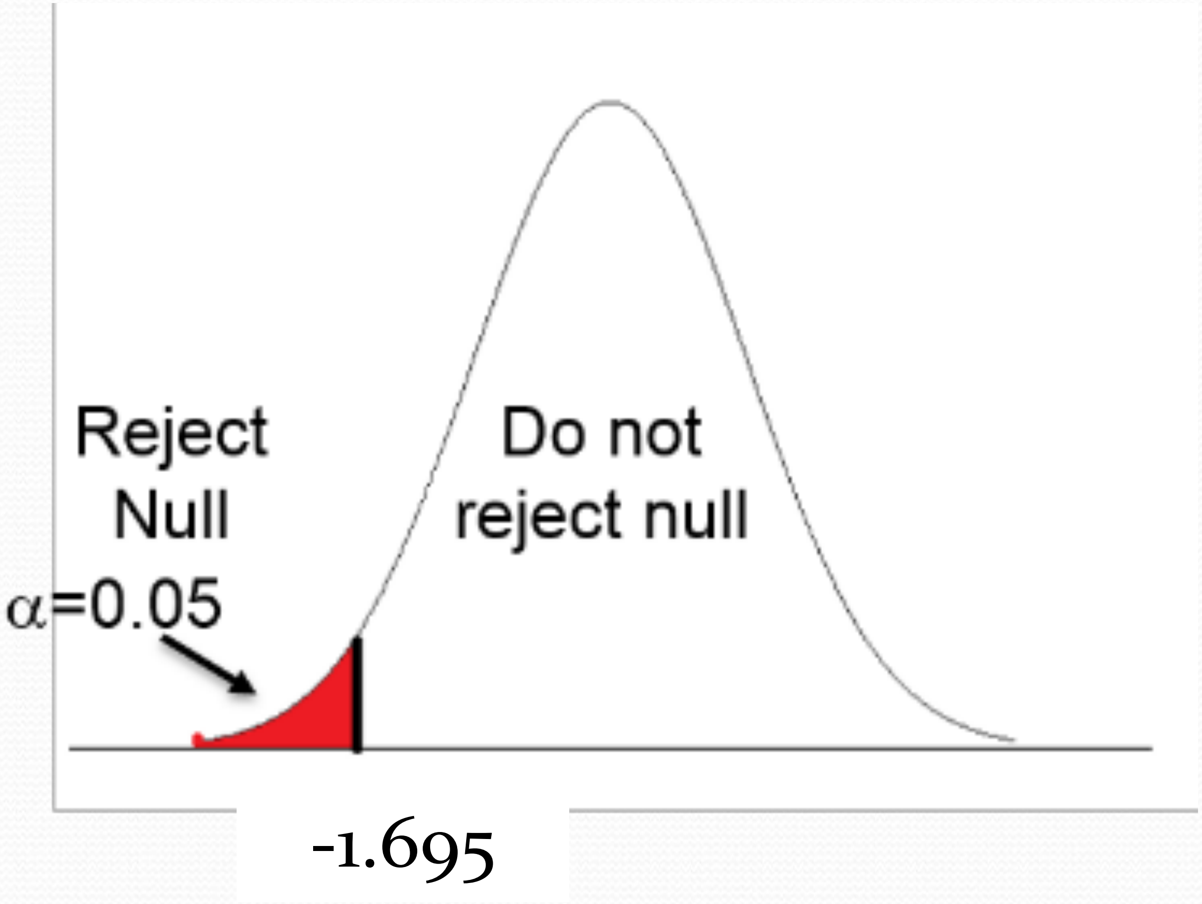
(3) Determine the P value and make a conclusion

Because $t_{0.05,35}$ =1.690, So $p<0.05$

Conclusion: **Reject H_o and accept H_1, the mean of hemoglobin concentrations of these workers is smaller than that of the healthy individuals (140g/L).**

t-test table

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$	$t_{.9995}$
one-tail	**0.50**	**0.25**	**0.20**	**0.15**	**0.10**	**0.05**	**0.025**	**0.01**	**0.005**	**0.001**	**0.0005**
two-tails	**1.00**	**0.50**	**0.40**	**0.30**	**0.20**	**0.10**	**0.05**	**0.02**	**0.01**	**0.002**	**0.001**
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.000	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.000	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.000	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.000	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.000	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.000	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.000	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.000	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.000	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.000	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	0.000	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	0.000	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	3.232	3.460
80	0.000	0.678	0.846	1.043	1.292	1.664	1.990	2.374	2.639	3.195	3.416
100	0.000	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	3.174	3.390
1000	0.000	0.675	0.842	1.037	1.282	1.646	1.962	2.330	2.581	3.098	3.300
z	0.000	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.090	3.291
	0%	50%	60%	70%	80%	90%	95%	98%	99%	99.8%	99.9%
	Confidence Level										


Reject
Null
α=0.05
Do not
reject null
-1.695

2. Paired *t* test

2. Paired t test

Paired *t*-test is suitable for comparing two sample means of a paired design.

Can you give any examples of a paired design?

Testing condition:

The testing condition of paired *t*-test includes the assumption of normal distribution of the differences.

The test statistic of a paired *t* test is

$$t = \frac{\bar{d} - 0}{S_d / \sqrt{n}} \qquad \nu = n - 1$$

Where $\bar{d}$ is the sample mean of paired measurement differences, S_d is the standard deviation of the differences, *n* is the number of the pairs and *v* is the degree of freedom.

Example:

We collected the systolic blood pressure levels in 10 women before and after using oral contraceptive (Table 1).

Question??

Is there any difference of systolic blood pressure levels in women before and after using oral contraceptive ?

Table 1 Systolic blood pressure levels in 10 women before and after using oral contraceptive

i	Before	After	d_i
1	115	128	13
2	112	115	3
3	107	106	-1
4	119	128	9
5	115	122	7
6	138	145	7
7	126	132	6
8	105	109	4
9	104	102	-2
10	115	117	2

We find that differences d_i obey normal distribution:

(1) Establish the hypothesis and determine the significance level

(a) H_0:

μ_d=0, there is no difference of systolic blood pressure levels in women before and after using oral contraceptive.

(b) H_1:

$\mu_d \neq 0$, the systolic blood pressure levels in women before using oral contraceptive are different from those after using oral contraceptive .

$\alpha = 0.05$

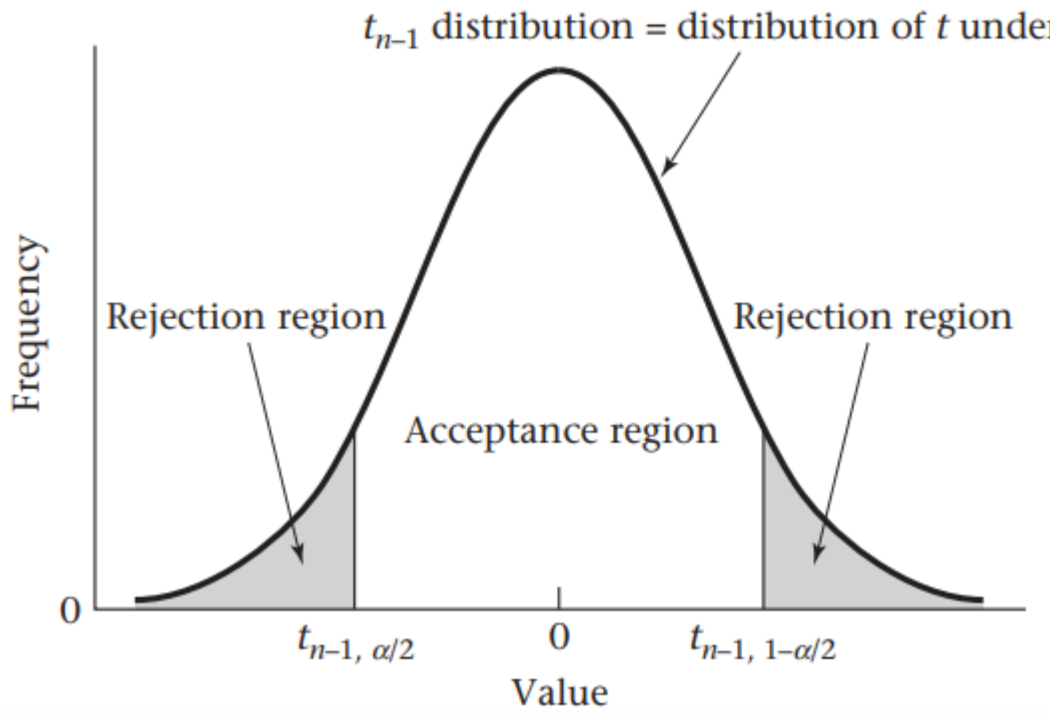
(2) Calculate the *t* test statistic

$$t = \frac{\bar{d}}{S/\sqrt{n}} = 3.32 \qquad \nu = n - 1 = 10 - 1 = 9$$

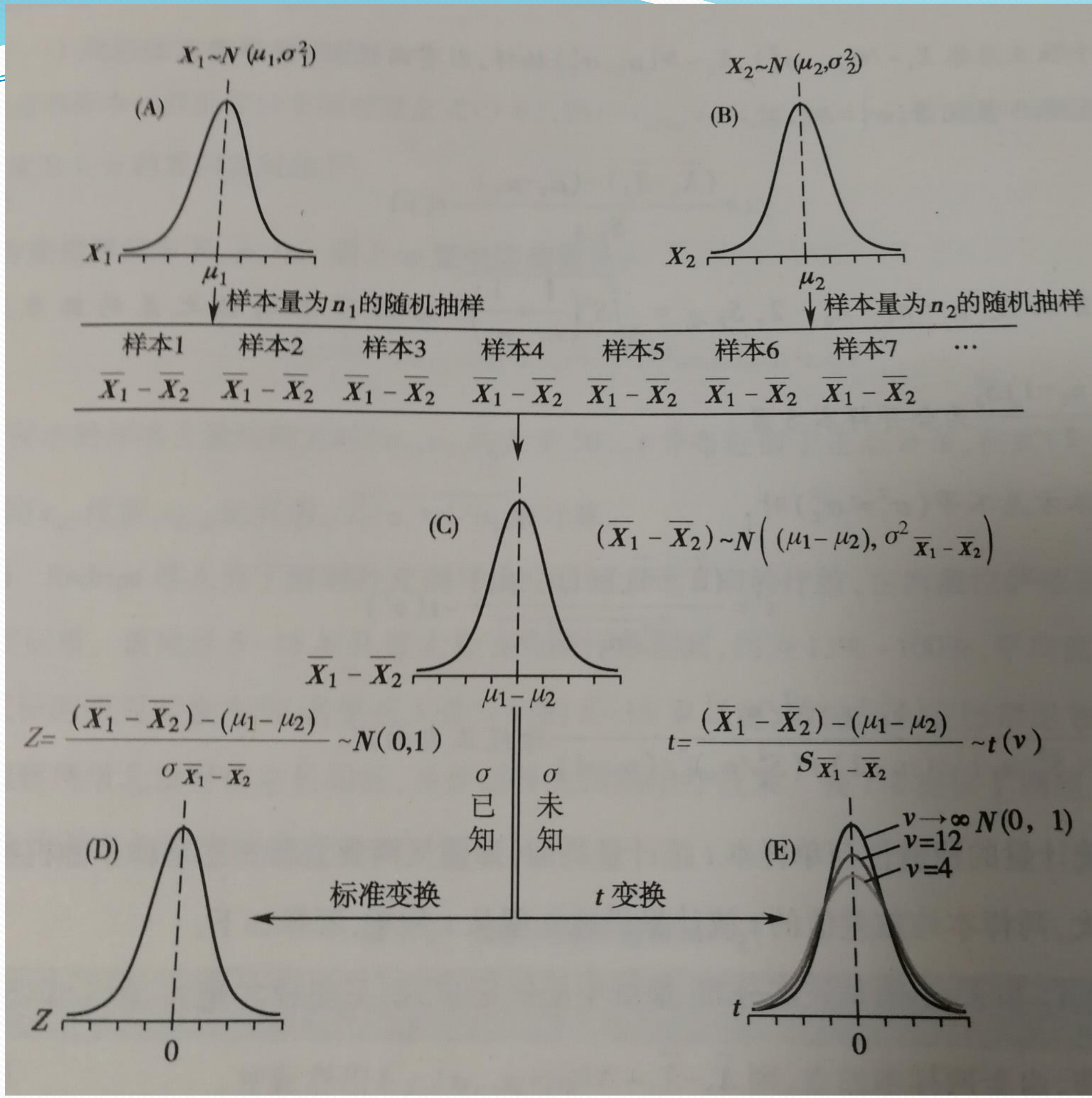
(3) Determine the *P* value and make a conclusion

Because $t_{0.05/2,\ 9} = 2.262$, So $p < 0.05$

Conclusion: **Reject H_o and accept H_1, the systolic blood pressure levels in women before using oral contraceptive are different from those after using oral contraceptive .**


t_{n-1} distribution = distribution of t under H_0
Frequency
Rejection region
Rejection region
Acceptance region
0
$t_{n-1,\ \alpha/2}$
0
$t_{n-1,\ 1-\alpha/2}$
Value

3. *t* test for two independent samples



X1~N (μ1,σ1²)
(A)
X1
μ1
样本量为n1的随机抽样
X2~N (μ2,σ2²)
(B)
X2
μ2
样本量为n2的随机抽样
样本1 样本2 样本3 样本4 样本5 样本6 样本7 …
X̄1 - X̄2
(C)
(X̄1 - X̄2)~N((μ1-μ2), σ²X̄1-X̄2)
X̄1 - X̄2
μ1-μ2
Z= ((X̄1 - X̄2)-(μ1-μ2)) / σX̄1-X̄2 ~N(0,1)
σ已知
σ未知
t= ((X̄1 - X̄2)-(μ1-μ2)) / SX̄1-X̄2 ~t(ν)
标准变换
t 变换
(D)
Z
0
(E)
ν→∞ N(0, 1)
ν=12
ν=4
t
0

Equal variance

$$s_{\bar{x}_1-\bar{x}_2}=\sqrt{s_C^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$$

$(n_1-1)/(n_1+n_2-2)$ W_1

$(n_2-1)/(n_1+n_2-2)$ W_2

$$s_C^2 = W_1 \times s_1^2 + W_2 \times s_2^2$$

$$\Downarrow$$

$$s_C^2=\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$$

3. *t* test for two independent samples

t test for two independent samples is suitable for comparing two sample means of a completely randomized design, and its purpose is to infer whether two population means are the same.

The testing conditions include the assumptions of

- ✓normal distribution
- ✓equal variances

The test statistic for independent two-sample *t* test is:

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - 0}{S_{\bar{X}_1 - \bar{X}_2}} = \frac{\bar{X}_1 - \bar{X}_2}{S_{\bar{X}_1 - \bar{X}_2}} \qquad v = n_1 + n_2 - 2$$

$$Where \quad S_{\bar{X}_1 - \bar{X}_2} = \sqrt{S_c^{\,2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

$$When \quad S_c^{\,2} = \frac{(n_1 - 1)S_1^{\,2} + (n_2 - 1)S_2^{\,2}}{n_1 + n_2 - 2}$$

where n_1 and n_2 are the sample sizes of the two groups respectively, and S_c^2 is the pooled variance of the two groups.

If the variances of the two populations are not equal, **approximate *t* test** is recommended.

- **Welch's *t* test**

$$t' = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^{\,2}}{n_1} + \frac{S_2^{\,2}}{n_2}}} \qquad v = \frac{(S_1^2/n_1 + S_2^2/n_2)^2}{\frac{(S_1^2/n_1)^2}{n_1+1} + \frac{(S_2^2/n_2)^2}{n_2+1}} - 2$$

- **Scatterthwaite's test**

$$t' = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^{\,2}}{n_1} + \frac{S_2^{\,2}}{n_2}}} \qquad v = \frac{(S_1^2/n_1 + S_2^2/n_2)^2}{\frac{(S_1^2/n_1)^2}{n_1-1} + \frac{(S_2^2/n_2)^2}{n_2-1}} - 2$$

Example:

To compare the effects of two drugs on the blood sugar reduction, 40 volunteer patients were randomly assigned to two groups receiving two drugs, respectively. Data of blood sugar reduction were collected as follows. Are the two drugs different in terms of lowering blood sugar?

Drug 1:

-0.70 -5.60 2.00 2.80 0.70 3.50 4.00 5.80 7.10 -0.50 2.50 -1.60 1.70 3.00 0.40 4.50 4.60 2.50 6.00 -1.40

Drug 2:

3.70 6.50 5.00 5.20 0.80 0.20 0.60 3.40 6.60 -1.10 6.00 3.80 2.00 1.60 2.00 2.20 1.20 3.10 1.70 -2.00

After the hypothesis test of normal distribution and equal variance, we conclude that our data meet the assumptions of normal distribution and equal variances.

(1) Establish the hypothesis and determine the significance level

(a) H_0:

$\mu_1=\mu_2$, the blood sugar reduction by drug 1 is equal to that by drug 2.

(b) H_1:

$\mu_1\neq\mu_2$, the blood sugar reduction by drug 1 is different from that by drug 2 .

$\alpha = 0.05$

(2) Calculate the *t* test statistic

$$t == \frac{\bar{X}_1 - \bar{X}_2}{S_{\bar{X}_1 - \bar{X}_2}} = -0.642 \quad v = n_1 + n_2 - 2 = 38$$

(3) Determine the *P* value and make a conclusion

Because $t_{0.05/2,38}$ **=2.024,** So **p>0.05**

Conclusion: Can not **reject *H_o*. We cannot conclude that the blood sugar reduction by drug 1 is different from that by drug 2.**

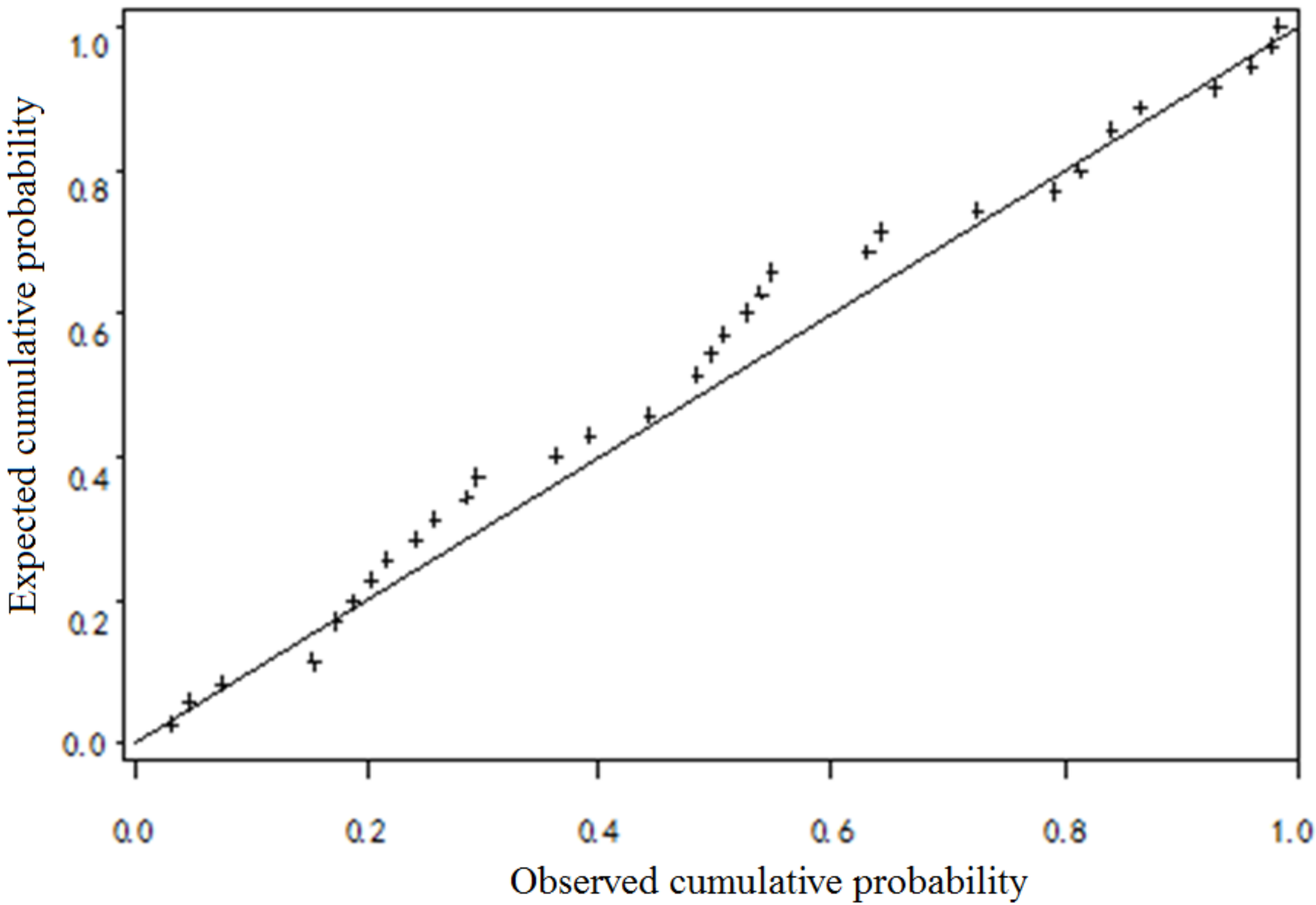
4. Tests for normality and equal variance

4.1 Normality test

(1) Graphical methods

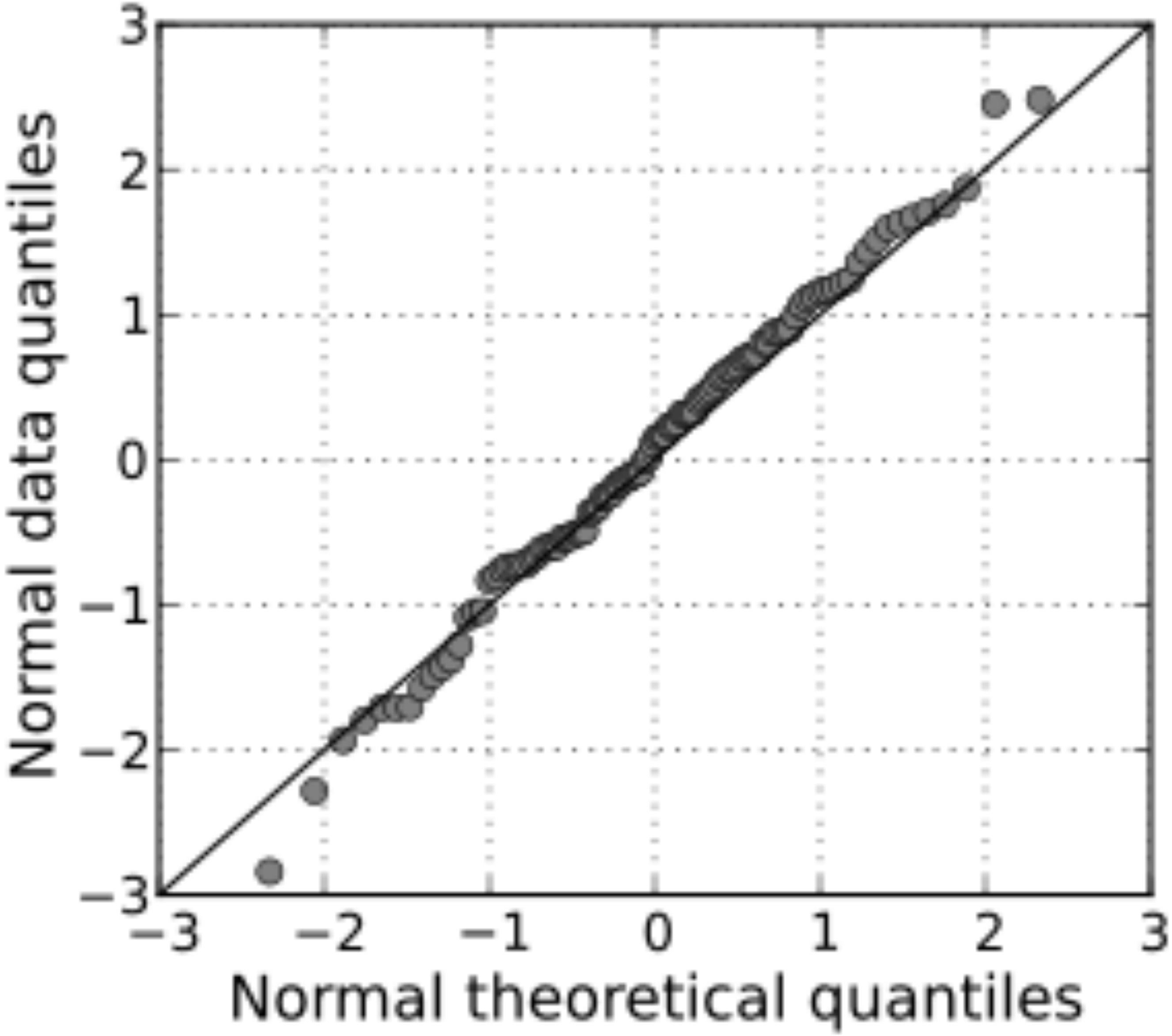
Probability-probability plot (P-P plot)

- **One axis: the cumulative probability of actual observation values**
- **Another axis: the expected/theoretical cumulative probability based on the normal distribution.**
- **A normal distribution means that sample points are distributed around the diagonal of the first quadrant.**


Expected cumulative probability
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
Observed cumulative probability

Quantile-quantile plot (Q-Q plot)

- **One axis: the quartile of sample data.**
- **Another axis: the expected/theoretical quartile based on the normal distribution.**
- **A normal distribution means that sample points are distributed around the diagonal of the first quadrant.**
- **Q-Q plot is more widely used than P-P plot in practice.**

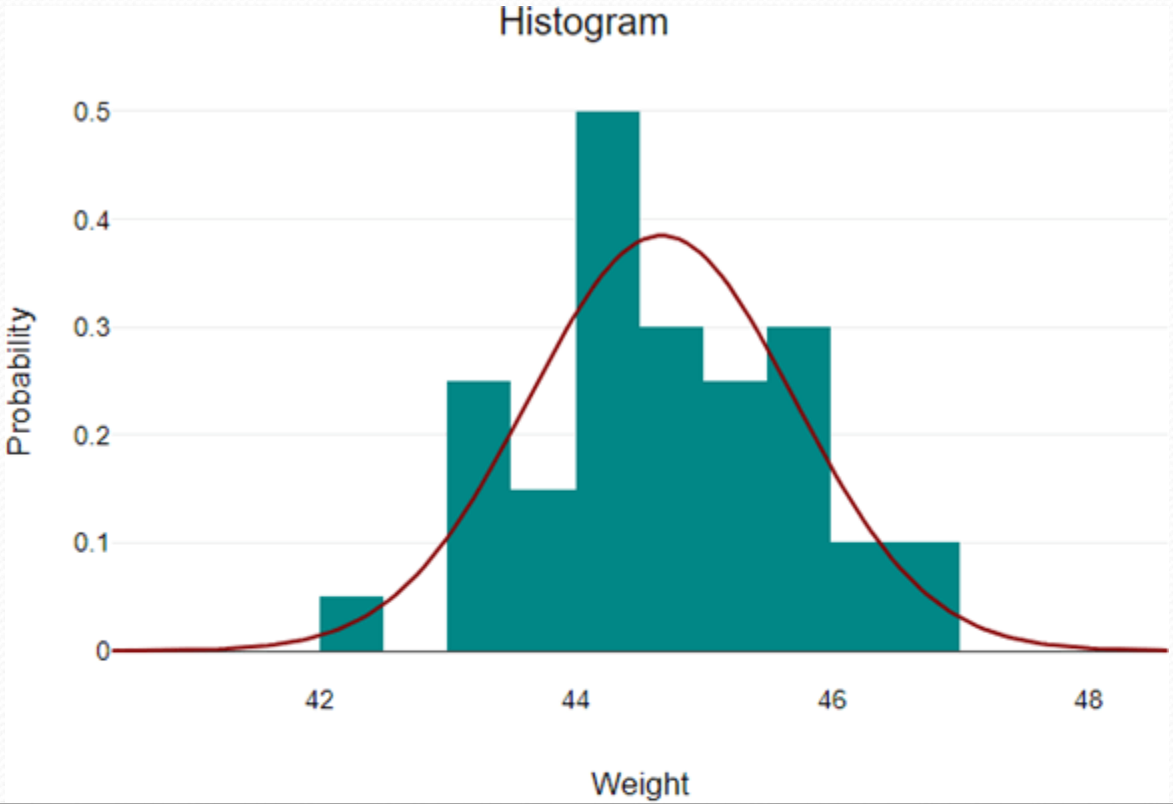


Normal data quantiles
Normal theoretical quantiles
3
2
1
0
−1
−2
−3
−3
−2
−1
0
1
2
3

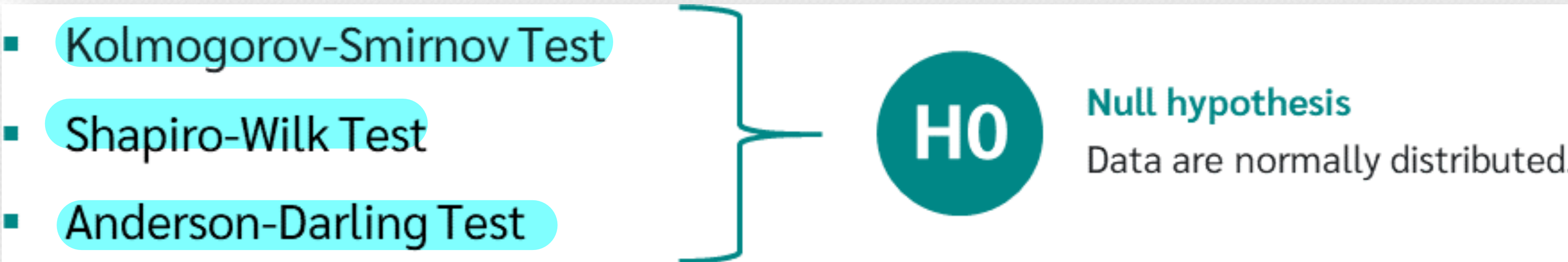
histogram

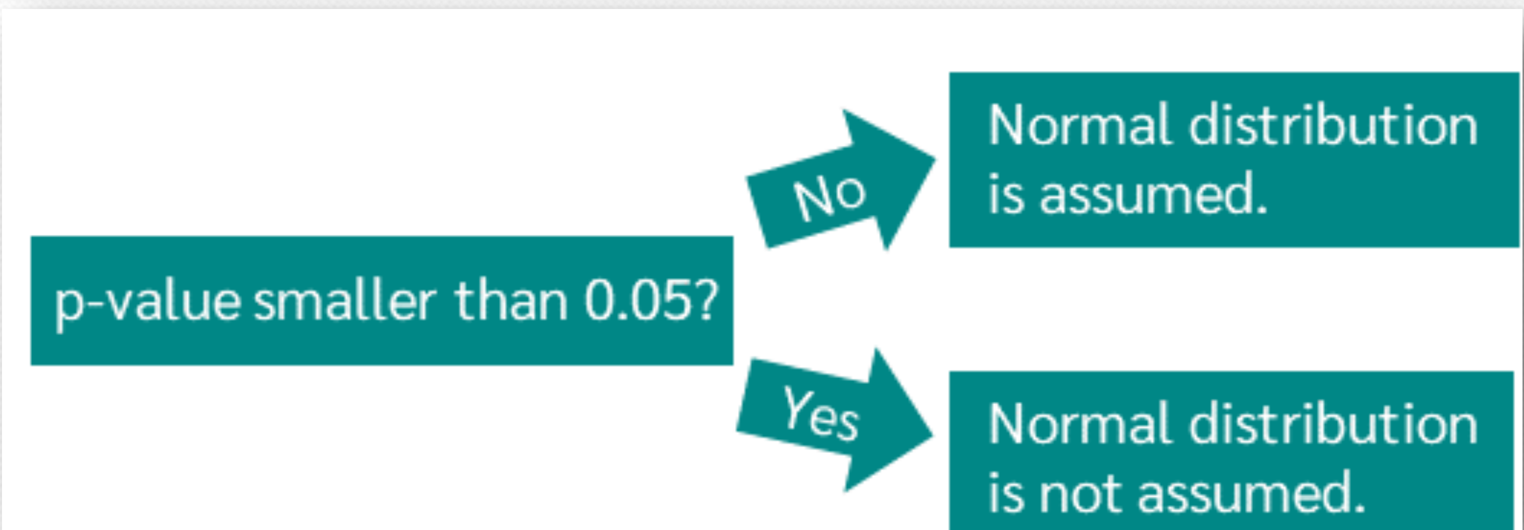
Y axis: the number of times that the values occurred within the intervals set by the X-axis.

- **X axis: intervals that show the scale of values which the measurements fall under.**
-



(2) Hypothesis testing methods





(2) Hypothesis testing methods

Shapiro-Wilk test

- **Known as W test and introduced by S.S.Shapiro and M.B.Wilk;**
- **Suitable for sample sizes in the range of 3 to 50;**
- **Shapiro-Wilk Expanded Test: a revised approach using the algorithm of J. P. Royston which can handle samples with up to 5,000 (or even more)**

Shapiro-Wilk Royston Test | Real Statistics Using Excel (real-statistics.com)

Basic approach of Shapiro-Wilk test

- Arrange the data in ascending order: x1≤x2 ≤x3….. ≤x4
- calculate SS $$SS = \sum_{i=1}^{n}(x_i - \bar{x})^2$$
- If n is even, let m = n/2, while if n is odd let m = (n–1)/2
- Calculate b as follows, taking the ai weights from Shapiro-Wilk Tables $$b = \sum_{i=1}^{m} a_i(x_{n+1-i} - x_i)$$
- W=b²/SS
- Find the value the Shapiro-Wilk Tables (for a given value of n) that is closest to W, interpolating if necessary.

$$W = \frac{\left[\sum_{i=1}^{n/2} a_i (x_{n+1-i} - x_i)\right]^2}{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}$$

Kolmogorov-Smirnov *test*

- **Suitable for sample sizes in the range of 50 to 1000;**
- **The formula for the test statistic is:**

$$Y = \frac{\sqrt{n}(D - 0.28209479)}{0.02998598}$$

in which

$$D = \frac{\sum_{i=1}^{n}\left(i - \frac{n+1}{2}\right)x_i}{(\sqrt{n})^3 \sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}}$$

[Kolmogorov-Smirnov Normality | Real Statistics Using Excel (real-statistics.com)](real-statistics.com)

4.1 Equal variances test

(1) F test

Testing whether two population variances are equal

$$F = \frac{s_1^2\left(bigger\right)}{s_2^2\left(smaller\right)}$$

$$\nu_1 = n_1 - 1, \quad \nu_2 = n_2 - 1$$

- ***Sensitive to departures from normality***

(2) Levene’s test

- Testing whether two or more population variances are equal
- less sensitive to departures from normality

(3) Bartlett's test

- Testing whether two or more population variances are equal
- Sensitive to departures from normality

Summary

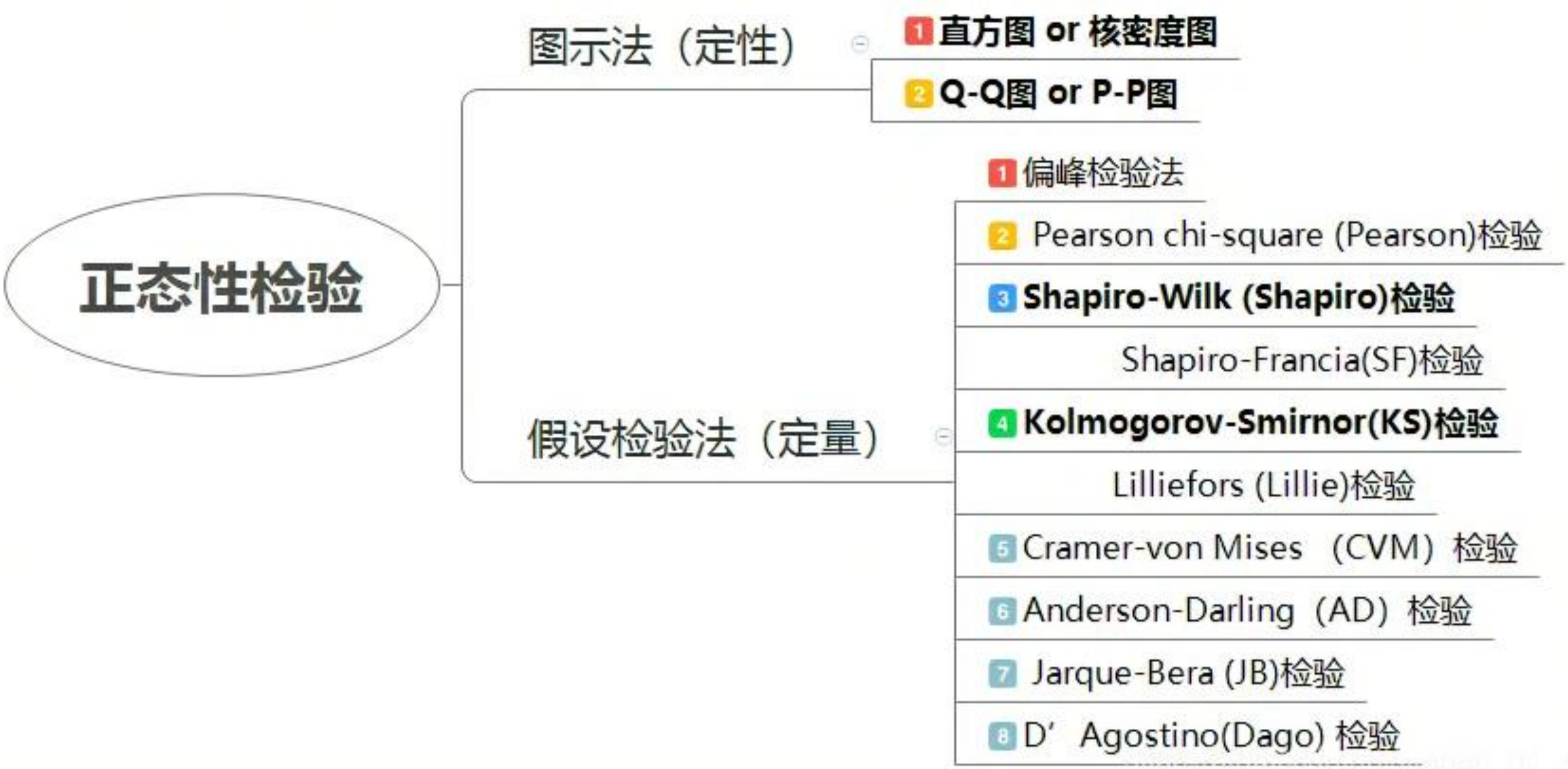
✓ **The *t* test, developed by Gosset, is one common hypothesis test for comparing two means.**

✓ **There are three types of *t* test: <u>one sample *t* test</u>, <u>paired *t* test</u> and <u>*t* test for two independent samples.</u>**

✓ **The testing conditions include the assumptions of**

- **normal distribution**
- **equal variances**

Thank you!


正态性检验
图示法（定性）
1 直方图 or 核密度图
2 Q-Q图 or P-P图
假设检验法（定量）
1 偏峰检验法
2 Pearson chi-square (Pearson)检验
3 Shapiro-Wilk (Shapiro)检验
Shapiro-Francia(SF)检验
4 Kolmogorov-Smirnor(KS)检验
Lilliefors (Lillie)检验
5 Cramer-von Mises （CVM）检验
6 Anderson-Darling （AD）检验
7 Jarque-Bera (JB)检验
8 D’ Agostino(Dago) 检验