第4-4章 假设检验实例

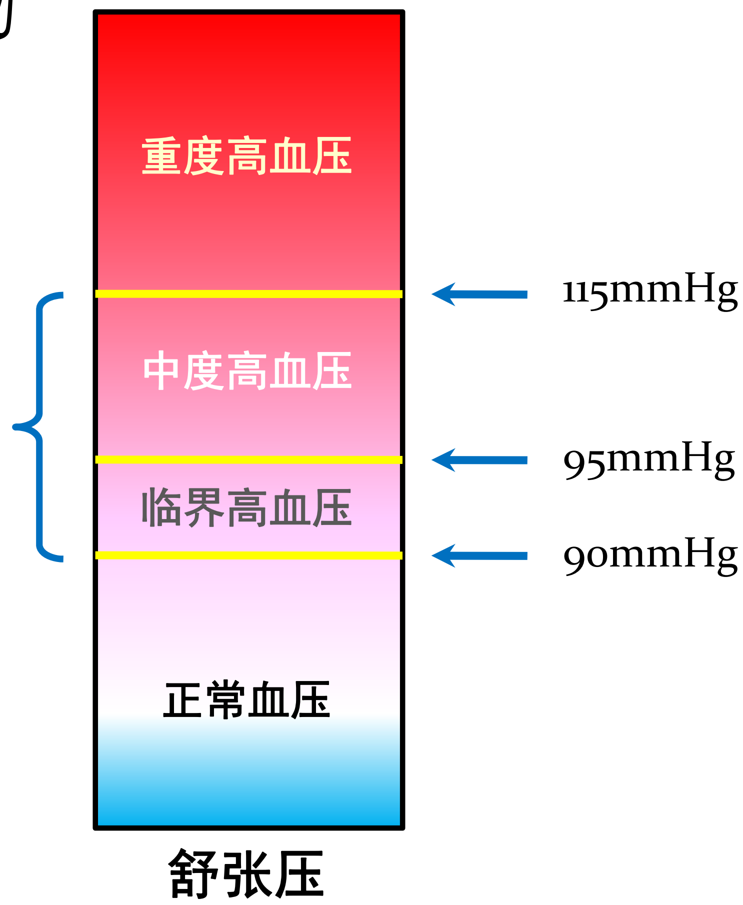
感谢清华大学自动化系江瑞教授提供PPT

内容

- 参数检验
  - 二项精确检验
  - Fisher精确检验
  - 正态分布均值和方差检验
- 非参数检验
  - 位置检验(符号检验，符号秩检验，秩和检验)
  - 拟合优度检验

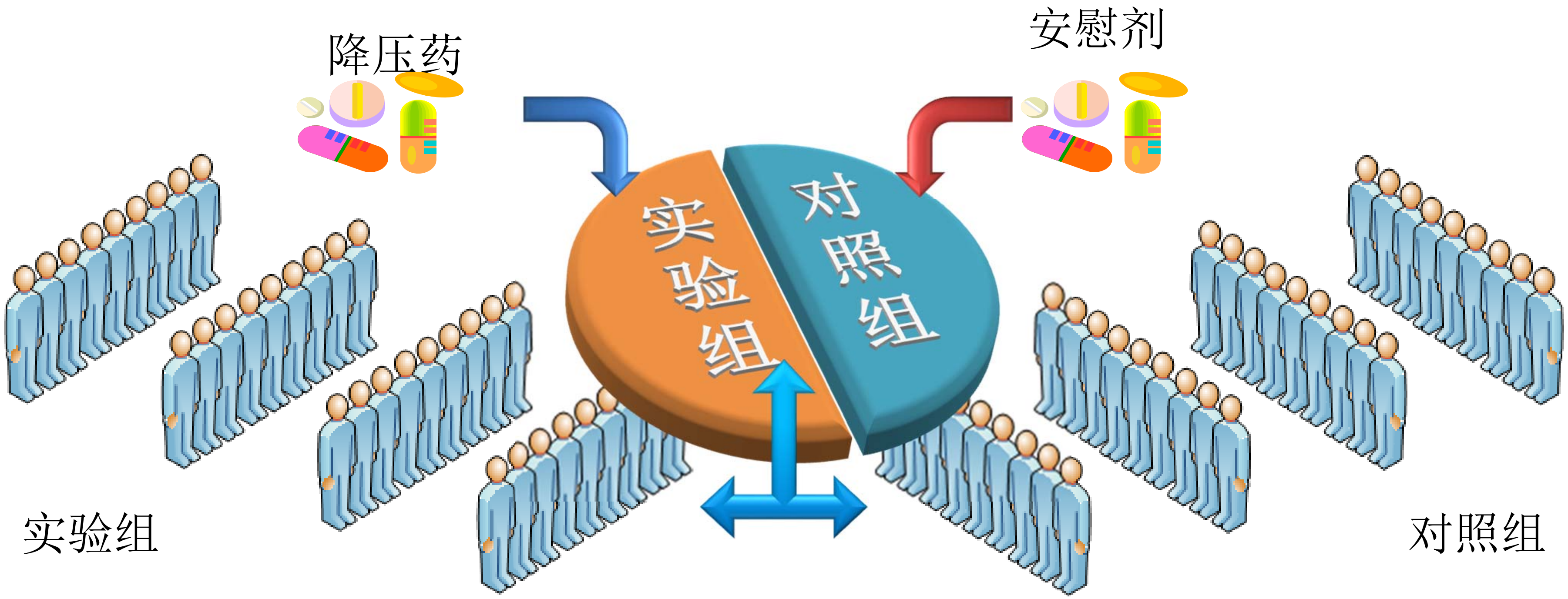
从实例谈起
降压药对中度及临界高血压患者的作用

- 某制药公司研制一种降压药物
- 对于重度高血压患者
  - 该药物的降压作用已经清楚了
- 对于中度和临界高血压患者
  - 该药物的降压作用还不清楚
- 通过实验研究降压药物对于中度和临界高血压患者(90-115 mmHg)的降压作用



实验设计

- 从**中度和临界高血压患者**中**随机抽取389**名，分为两组
- 实验组**193**名患者，服用降压药
- 对照组**196**名患者，服用对照剂



原始数据 — 实验组初始血压

100	99	107	96	104	96	94	105	97	91
105	114	93	98	102	98	104	102	100	93
113	99	99	97	102	102	105	114	110	109
115	91	94	112	95	102	95	91	104	111
101	107	114	101	102	99	108	91	94	110
106	114	93	95	95	90	94	96	94	112
110	96	114	97	99	115	106	103	103	106
94	93	92	90	104	113	102	93	95	92
114	105	97	93	95	95	102	115	104	104
108	90	94	109	109	95	95	105	111	109
90	114	94	96	98	105	114	100	113	115
115	92	99	103	93	99	102	114	102	96
98	109	96	112	115	98	109	96	105	106
92	93	93	91	100	114	106	115	96	95
113	99	110	110	104	114	102	92	92	95
108	110	101	99	113	111	111	103	100	91
96	94	91	108	102	93	90	93	109	108
108	114	111	90	104	100	90	95	109	101
98	113	103	96	110	96	97	96	94	91
97	113	110							

原始数据—实验组12个月后血压

77	88	100	83	108	64	74	87	104	59
91	96	99	108	98	83	83	92	110	73
98	89	88	87	102	89	91	108	119	90
105	72	88	117	78	96	80	97	96	92
85	109	108	87	80	88	85	85	82	98
106	114	90	81	61	84	64	79	77	111
80	65	97	89	107	92	94	112	70	110
76	77	64	80	96	93	94	73	87	111
102	94	78	81	85	75	91	113	87	80
90	77	119	113	96	94	88	103	89	84
71	106	77	86	71	104	108	81	117	99
86	96	91	87	77	100	95	83	93	96
94	93	91	106	101	95	109	127	95	83
78	69	82	89	120	109	85	114	94	92
92	94	100	98	104	109	112	66	83	85
124	95	75	100	95	92	95	113	100	95
87	90	92	101	97	82	82	88	90	103
106	114	93	57	76	77	72	82	101	98
96	114	109	76	99	90	95	64	67	69
80	101	103							

原始数据 — 对照组初始血压

97	105	110	103	90	94	115	111	114	99
105	113	110	103	90	106	93	93	91	113
113	91	100	99	104	96	114	98	101	92
106	106	95	94	98	98	109	93	112	104
105	91	113	111	115	109	98	108	114	115
103	102	113	113	104	110	112	97	112	98
103	99	100	104	104	115	99	103	113	107
97	96	107	115	114	102	103	96	93	94
101	90	91	107	100	109	92	90	112	98
99	108	97	97	113	106	91	96	91	100
110	109	105	96	115	113	107	109	96	102
92	96	113	113	112	100	104	97	101	115
110	109	103	115	94	102	94	94	94	111
99	110	112	109	95	98	107	93	111	96
105	114	99	91	111	102	105	91	104	111
113	92	102	91	112	114	101	107	112	94
95	110	105	97	91	106	112	94	99	110
93	91	110	101	109	115	114	108	111	94
109	97	112	115	113	110	105	114	115	90
92	104	109	104	115	90				

原始数据 — 对照组12个月后血压

82	103	116	94	87	93	124	126	131	102
115	103	92	105	105	103	92	103	96	133
99	85	103	109	101	97	130	98	101	87
112	92	96	102	89	108	115	83	116	101
93	96	130	113	135	112	90	92	102	102
97	107	130	121	99	102	103	109	105	77
93	97	96	86	110	107	91	113	133	112
86	77	94	134	108	92	101	104	95	81
112	98	91	90	100	93	69	110	91	92
103	103	85	80	93	100	93	91	96	102
110	124	106	100	133	128	126	92	91	92
78	104	117	133	111	110	116	92	106	110
130	116	110	111	94	100	95	94	95	111
99	110	102	116	99	98	107	67	113	102
125	137	97	102	107	95	125	95	107	131
136	90	113	87	105	134	105	110	132	109
97	100	107	81	90	100	115	75	106	116
100	93	132	105	103	135	79	105	134	87
106	102	122	130	105	142	132	136	132	102
99	106	106	96	102	72				

数据组织 — 实验组血压变化量

-23	-11	-7	-13	4	-32	-20	-18	7	-32
-14	-18	6	10	-4	-15	-21	-10	10	-20
-15	-10	-11	-10	0	-13	-14	-6	9	-19
-10	-19	-6	5	-17	-6	-15	6	-8	-19
-16	2	-6	-14	-22	-11	-23	-6	-12	-12
0	0	-3	-14	-34	-6	-30	-17	-17	-1
-30	-31	-17	-8	8	-23	-12	9	-33	4
-18	-16	-28	-10	-8	-20	-8	-20	-8	19
-12	-11	-19	-12	-10	-20	-11	-2	-17	-24
-18	-13	25	4	-13	-1	-7	-2	-22	-25
-19	-8	-17	-10	-27	-1	-6	-19	4	-16
-29	4	-8	-16	-16	1	-7	-31	-9	0
-4	-16	-5	-6	-14	-3	0	31	-10	-23
-14	-24	-11	-2	20	-5	-21	-1	-2	-3
-21	-5	-10	-12	0	-5	10	-26	-9	-10
16	-15	-26	1	-18	-19	-16	10	0	4
-9	-4	1	-7	-5	-11	-8	-5	-19	-5
-2	0	-18	-33	-28	-23	-18	-13	-8	-3
-2	1	6	-20	-11	-6	-2	-32	-27	-22
-17	-12	-7							

数据组织 — 对照组血压变化量

-15	-2	6	-9	-3	-1	9	15	17	3
10	-10	-18	2	-18	5	-3	-1	10	5
20	-14	-6	3	10	-3	1	16	0	0
-5	6	-14	1	8	-9	10	6	-10	4
-3	-12	5	17	2	20	3	-8	-16	-12
-13	-6	5	17	8	-5	-8	-9	12	-7
-21	-10	-2	-4	-18	6	-8	-8	10	20
5	-11	-19	-13	19	-6	-10	-2	8	2
-13	11	8	0	-17	0	-16	-23	20	-21
-6	4	-5	-12	-17	-20	-6	2	-5	5
2	0	15	1	4	18	15	19	-17	-5
-10	-14	8	4	20	-1	10	12	-5	5
-5	20	7	7	-4	0	-2	1	0	1
0	0	0	-10	7	4	0	0	-26	2
6	20	23	-2	11	-4	-7	20	4	3
20	23	-2	11	-4	-7	20	4	3	20
15	2	-10	2	-16	-1	-6	3	-19	7
6	7	2	22	4	-6	20	-35	-3	23
-7	-3	5	10	15	-8	32	27	22	17
12	7	2	-3	-8	-13				

检验问题

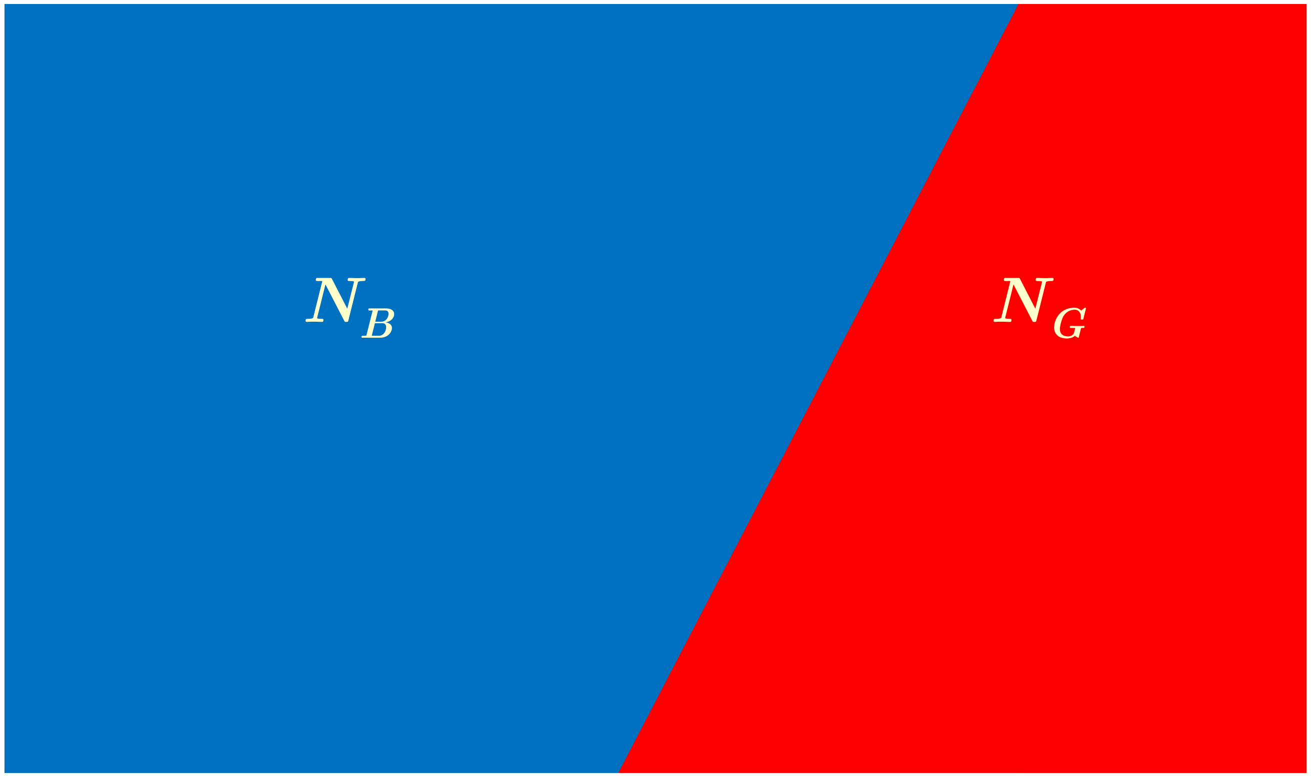
- 降压药物是否有降压作用？
- 降压药物的降压作用有多大？
- 降压药物是否能够阻止高血压的恶化？
- 降压药物是否有助于缓解高血压症状？
- 降压药物是否有助于缓解冠心病？

Binomial Exact Test

统计学方法及其应用 › 统计学基础 › 随机变量的函数

"A random variable is a quantity whose values are random and to which a probability distribution is assigned."

Is the proportion of girls in Tsinghua less than 50%



Tests of a single proportion

Protocol

A random sample $X_1, \dots X_n$ is observed from a Bernoulli population whose probability of success is p.

We like to test

(1) $H_0 : p = p_0$ versus $H_1 : p > p_0$;

(2) $H_0 : p = p_0$ versus $H_1 : p < p_0$;

(3) $H_0 : p = p_0$ versus $H_1 : p \neq p_0$.

Neyman-Pearson Test

Let $X_1,\dots,X_n \sim Bernoulli(\theta)$, iid. Consider the simple hypotheses

$$H_0 : \theta = \theta_0 \text{ versus } H_1 : \theta = \theta_1 \quad (\theta_0 < \theta_1).$$

ΣX is a sufficient statistic of θ, and the pmf of ΣX is

$$f(\Sigma x \mid \theta_i) \propto \theta_i^{\Sigma x}(1-\theta_i)^{n-\Sigma x},$$

Therefore

$$\frac{f(\Sigma x \mid \theta_1)}{f(\Sigma x \mid \theta_0)} = \frac{\theta_1^{\Sigma x}(1-\theta_1)^{n-\Sigma x}}{\theta_0^{\Sigma x}(1-\theta_0)^{n-\Sigma x}} = \left(\frac{\theta_1}{\theta_0}\frac{1-\theta_0}{1-\theta_1}\right)^{\Sigma x}\left(\frac{1-\theta_1}{1-\theta_0}\right)^{n},$$

Since $\theta_1 > \theta_0, 1-\theta_0 > 1-\theta_1$,

$$\frac{\theta_1}{\theta_0} > 1; \frac{1-\theta_0}{1-\theta_1} > 1; \frac{\theta_1}{\theta_0}\frac{1-\theta_0}{1-\theta_1} > 1$$

$f(\overline{x} \mid \mu_1) > kf(\overline{x} \mid \mu_0)$ is equivalent to

$$\Sigma x > \left[\log k + n\log\left(\frac{1-\theta_0}{1-\theta_1}\right)\right] \Big/ \log\left(\frac{\theta_1}{\theta_0}\frac{1-\theta_0}{1-\theta_1}\right) \equiv c(k).$$

Therefore, we like to reject H_0 when Σx is sufficiently large.

Karlin-Rubin Theorem

Let $X_1, \ldots, X_n \sim \text{Bernoulli}(\theta)$, iid. Consider the hypotheses

$$H_0 : \theta = \theta_0 \text{ versus } H_1 : \theta > \theta_0.$$

ΣX is a sufficient statistic of θ. Moreover, ΣX has a Binomial distribution, which has a monotone likelihood ratio. According to the Karlin-Rubin theorem, the test reject H_0 when $\Sigma X > t_0$ is a UMP level α test, where

$$\alpha = P(B_{n,\theta_0} > t_0).$$

Here, $B_{n,\theta}$ is a Binomial(n,θ) random variable.

p-value

***p*-value**

Let $W(\mathbf{X})$ be a test statistic such that **large** values of W give evidence that H_1 is true. For each sample point $\mathbf{x}$, define

$$p(\mathbf{x}) = \sup_{\theta\in\Theta_0} P_\theta(W(\mathbf{X}) \geq W(\mathbf{x})).$$

Then, $p(\mathbf{X})$ is a valid p-value.

Let $W(\mathbf{X})$ be a test statistic such that **small** values of W give evidence that H_1 is true. For each sample point $\mathbf{x}$, define

$$p(\mathbf{x}) = \sup_{\theta\in\Theta_0} P_\theta(W(\mathbf{X}) \leq W(\mathbf{x})).$$

Then, $p(\mathbf{X})$ is a valid p-value.

Binomial test, greater

One-sample binomial test, greater

Hypothesis:

$$H_0 : p = p_0 \quad \text{versus} \quad H_1 : p > p_0$$

Test statistic:

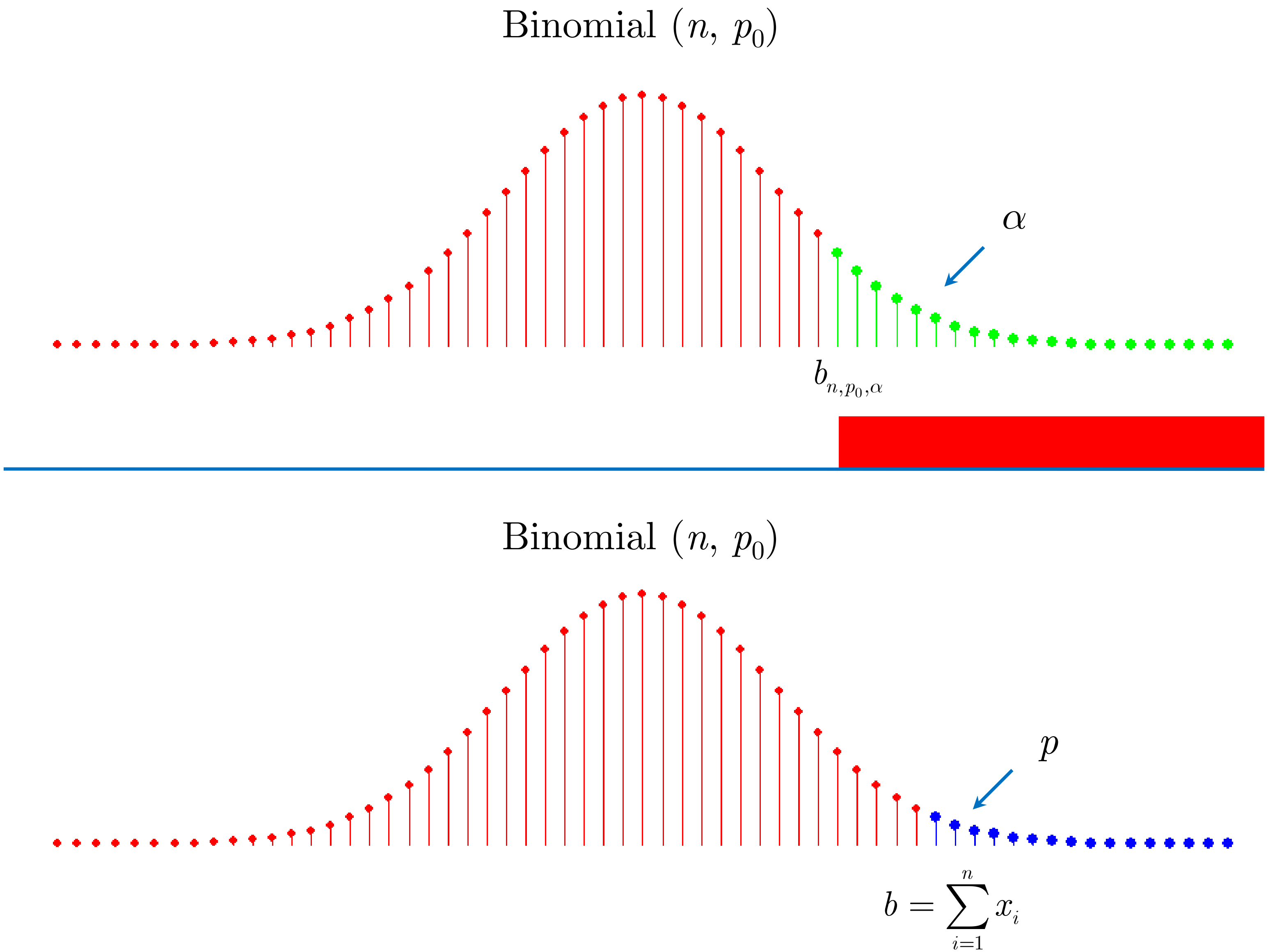
$$B = \sum_{i=1}^{n} X_i,$$

Large values of B give evidence that H_1 is true.

Under the null, B has a $Binomial(n,\ p_0)$ distribution.

Therefore, the p-value is

$$p(\mathbf{x}) = P\left(B_{n,p_0} \geq \sum_{i=1}^{n} x_i\right)$$



Binomial $(n,\ p_0)$
α
$b_{n,p_0,\alpha}$
Binomial $(n,\ p_0)$
p
$b = \sum_{i=1}^{n} x_i$

Binomial test, less

One-sample binomial test, less

Hypothesis:

$$H_0 : p = p_0 \quad \text{versus} \quad H_1 : p < p_0$$

Test statistic:

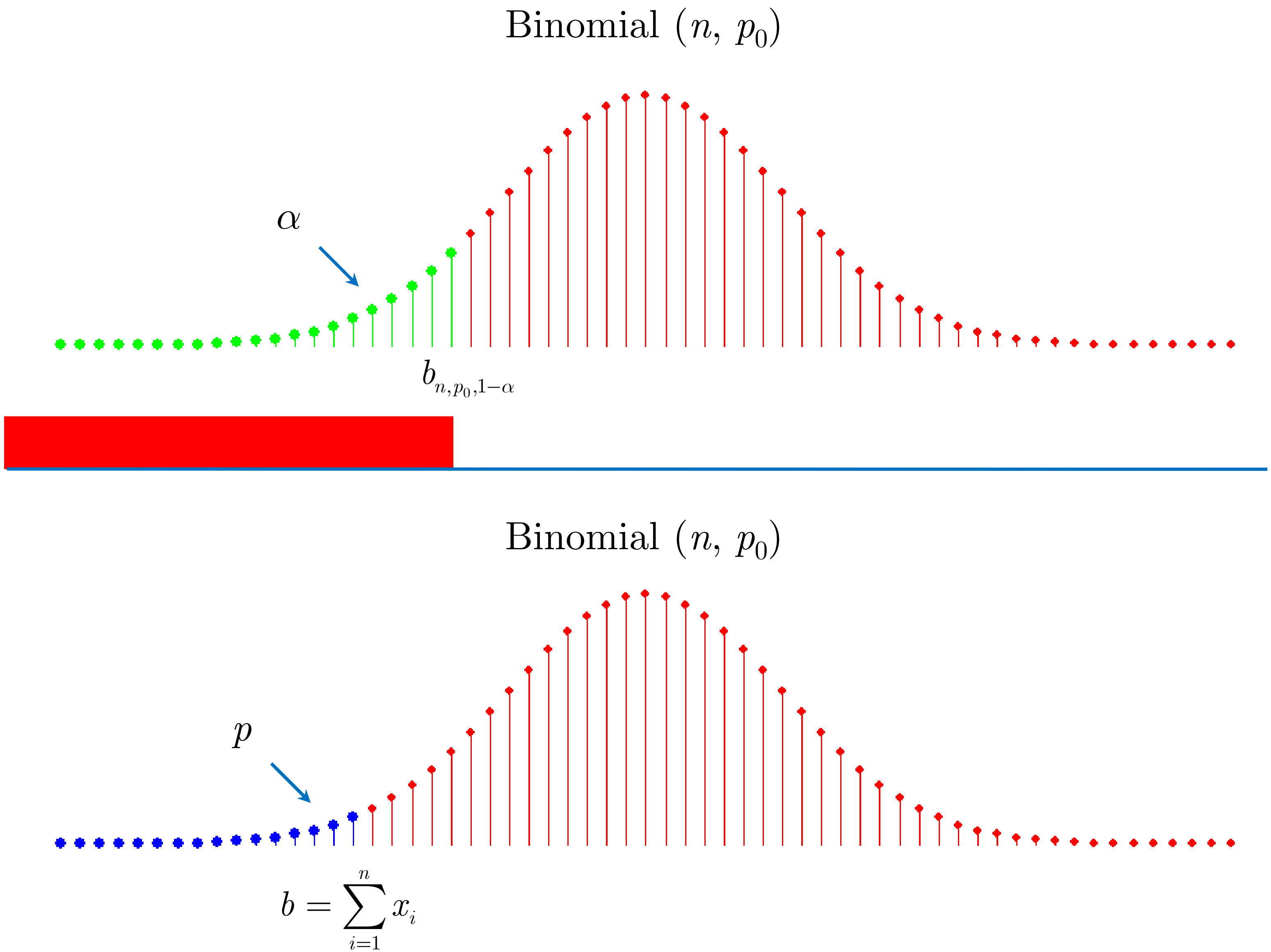
$$B = \sum_{i=1}^{n} X_i,$$

Small values of B give evidence that H_1 is true.

Under the null, B has a $Binomial(n,\ p_0)$ distribution.

Therefore, the p-value is

$$p(\mathbf{x}) = P\left(B_{n,p_0} \le \sum_{i=1}^{n} x_i\right)$$


Binomial $(n,\ p_0)$
α
$b_{n,p_0,1-\alpha}$
Binomial $(n,\ p_0)$
p
$b = \sum_{i=1}^{n} x_i$

Binomial test, two-sided

One-sample binomial test, two-sided

Hypothesis:

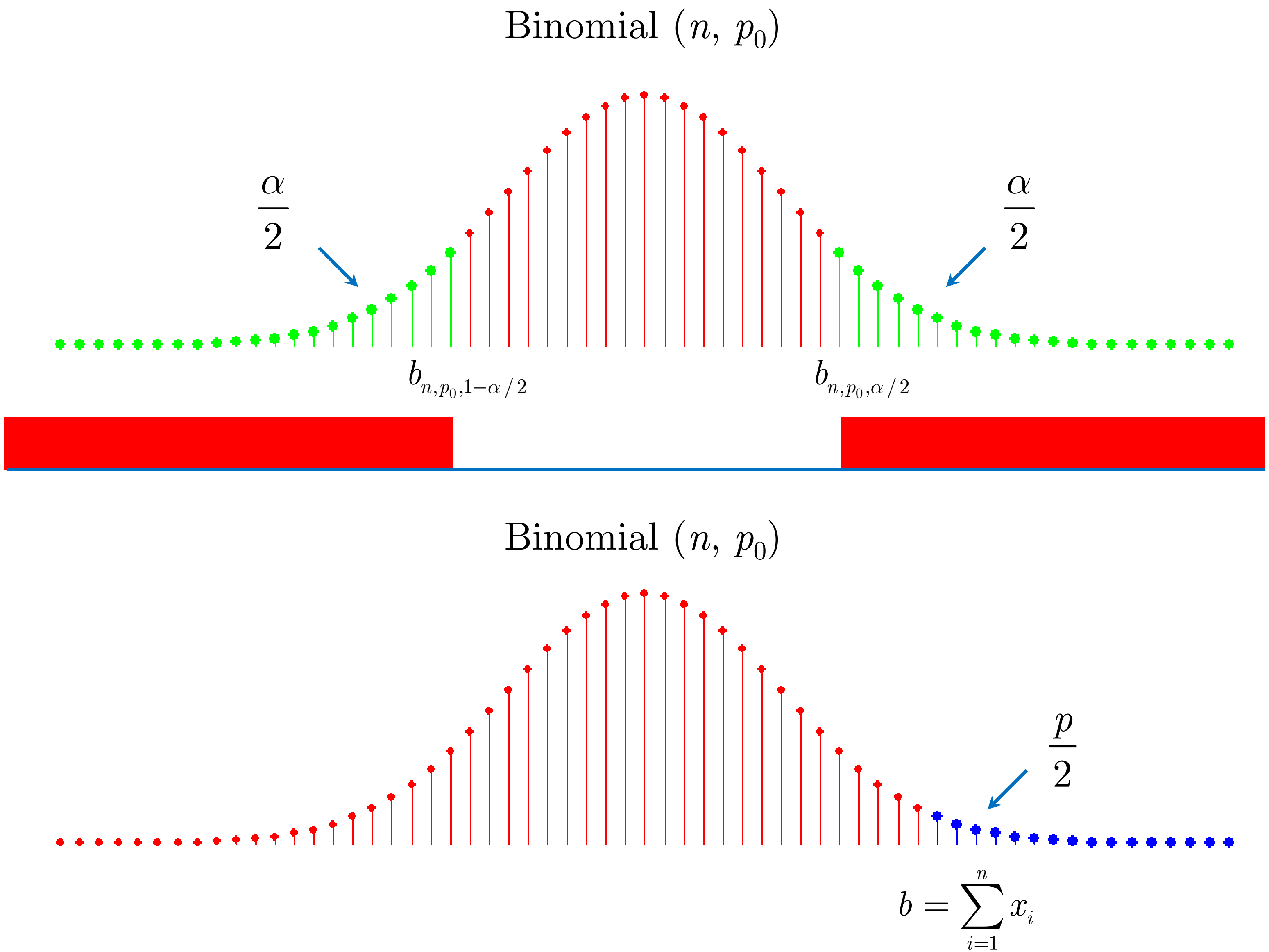
$$H_0 : p = p_0 \quad \text{versus} \quad H_1 : p \neq p_0$$

Test statistic:

$$B = \sum_{i=1}^{n} X_i,$$

p-value:

$$p(\mathbf{x}) = \begin{cases} 2P\left(B_{n,p_0} \geq \sum_{i=1}^{n} x_i\right), & \sum_{i=1}^{n} x_i \geq np_0 \\ 2P\left(B_{n,p_0} \leq \sum_{i=1}^{n} x_i\right), & \sum_{i=1}^{n} x_i < np_0 \end{cases}$$



Binomial $(n,\ p_0)$
$\frac{\alpha}{2}$
$\frac{\alpha}{2}$
$b_{n,p_0,1-\alpha/2}$
$b_{n,p_0,\alpha/2}$
Binomial $(n,\ p_0)$
$\frac{p}{2}$
$b = \sum_{i=1}^{n} x_i$

Practice

Test whether a coin is fair

```
> binom.test(520, 1000, p=0.5, alternative="two.sided")

         Exact binomial test

data:  520 and 1000
number of successes = 520, number of trials = 1000, p-value = 0.2174
alternative hypothesis: true probability of success is not equal to 0.5
...

> binom.test(540, 1000, p=0.5, alternative="two.sided")

        Exact binomial test

data:  540 and 1000
number of successes = 540, number of trials = 1000, p-value = 0.01244
alternative hypothesis: true probability of success is not equal to 0.5
...
```

Normal Approximation

- Let $X_1, \cdots, X_n \sim Bernoulli(p),$ consider the test
$$H_0 : p = p_0 \leftrightarrow H_1 : p > p_0$$
- Since for Bernoulli population
$$\mu = EX = p, \sigma^2 = var(X) = p(1-p)$$
- From the central limit theorem and the Slutskey theorem,
$$Z_n = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} = \frac{\hat{p} - p}{\sqrt{\hat{p}(1-\hat{p})/n}} \to N(0,1)$$
- Where $\hat{p} = \overline{X}$ is the estimator of p, and $\hat{p}(1-\hat{p})$ is that of σ^2.
- This suggests a test with test statistic

$$Z_n = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} = \frac{\hat{p} - p}{\sqrt{\hat{p}(1-\hat{p})/n}}$$

With rejection region $\{Z_n > z_\alpha\}$

χ^2 approximation

Since the square of a standard normal random variable is a χ_1^2 random variable. We have that

$$Z_n^2 = \left(\frac{\bar{X} - \mu}{\sigma / \sqrt{n}}\right)^2 \Rightarrow \frac{(\hat{p} - p)^2}{\hat{p}(1 - \hat{p}) / n} \sim \chi_1^2.$$

This suggests the use of Z_n^2 as the test statistic and rejects H_0 if and only if $Z_n^2 > \chi_{1,\alpha}^2$.

Practice

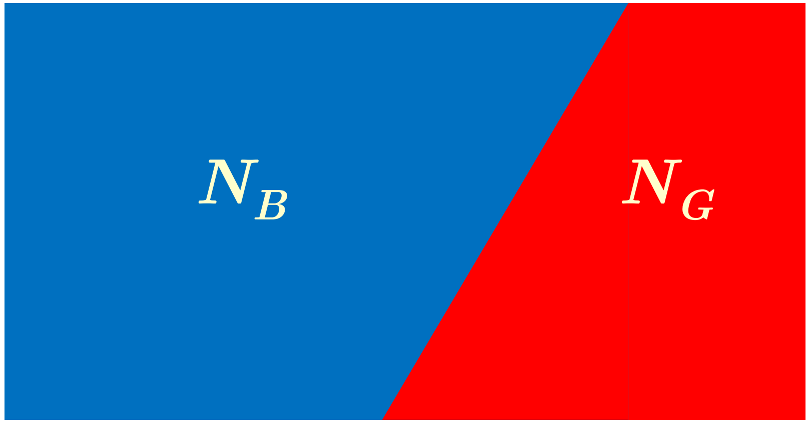
Test whether a coin is fair

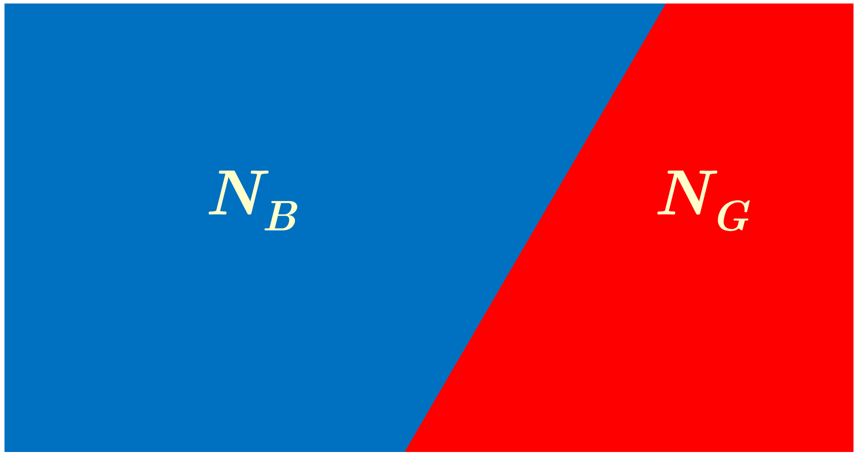
```
> prop.test(520, 1000, p=0.5, alternative="two.sided")

         1-sample proportions test with continuity correction

data:  520 out of 1000, null probability 0.5
X-squared = 1.521, df = 1, p-value = 0.2175
alternative hypothesis: true p is not equal to 0.5
...

> prop.test(540, 1000, p=0.5, alternative="two.sided")

         1-sample proportions test with continuity correction

data:  540 out of 1000, null probability 0.5
X-squared = 6.241, df = 1, p-value = 0.01248
alternative hypothesis: true p is not equal to 0.5
...
```

Fisher's Exact Test

"A random variable is a quantity whose values are random and to which a probability distribution is assigned."

Is the proportion of girls in Peking University significantly larger than that of Tsinghua?





Tests of two proportions

Protocol

Let Populatoin X and Population Y in a contingency have Bernoulli (p_X) and Bernoulli (p_Y) distributions, respectively. We like to test

(1) $H_0 : p_X = p_Y$ versus $H_1 : p_X > p_Y$;

(2) $H_0 : p_X = p_Y$ versus $H_1 : p_X < p_Y$;

(3) $H_0 : p_X = p_Y$ versus $H_1 : p_X \neq p_Y$.

Tests of two proportions

Contingency table

	Population X	Population Y	Total
Successes	s_x	s_y	$s = s_x + s_y$
Failures	f_x	f_y	$f = f_x + f_y$
Total	n_x	n_y	$n = n_x + n_y$

Conditional on a sufficient statistic

Conditional on a sufficient statistic to define a p-value

Let $W(\mathbf{X})$ be a test statistic such that **large** values of W give evidence that H_1 is true. Let $S(\mathbf{X})$ be a sufficient statistic for the parameter θ **under the null model**. For each sample point $\mathbf{x}$, define

$$p(\mathbf{x}) = P(W(\mathbf{X}) \geq W(\mathbf{x}) \mid S = S(\mathbf{x})).$$

Then, $p(\mathbf{X})$ is a valid p-value.

Fisher's exact test

Since $S_X = \sum_{i=1}^{n_X} X_i$ is a sufficient statistic for p_X and $S_Y = \sum_{i=1}^{n_X} Y_i$ is a sufficient statistic for p_Y, we can make the decision with the use of only S_X and S_Y.

Obviously, S_X has a binomial $(n_X,\ p_X)$ distribution and S_Y has a binomial $(n_Y,\ p_Y)$ distribution. When **the null is true** $(p_X = p_Y = p)$, the joint pmf of (S_X, S_Y) is

$$f(s_X, s_Y \mid n_X, n_Y, p) = \binom{n_X}{s_X}\binom{n_Y}{s_Y} p^{s_X + s_Y}(1-p)^{(n_X + n_Y)-(s_X + s_Y)},$$

which suggests that $S = S_X + S_Y$ is a sufficient statistic for p under the null hypothesis.

Fisher's exact test

$$H_0 : p_X = p_Y \quad \text{versus} \quad H_1 : p_X > p_Y$$

When **the alternative is true** and $s = s_X + s_Y$ is fixed, we expect to see a large s_X and a small s_Y. So the count S_X can be used as a test statistic for p_X.

Now, we have the familiar thing: n_X red balls and n_Y white balls are mixed in an urn, randomly pick up s of them, what is the probability of observing exactly s_X red balls?

$$f(s_X \mid n_X, n_Y, s) = \binom{n_X}{s_X}\binom{n_Y}{s - s_X} \Big/ \binom{n_X + n_Y}{s},$$

that is, S_X has a hypergeometric distribution.

Fisher's exact test

Further, what is the probability of observing S_X values that are at least as extreme as s_X under the null?

$$P(S_X \geq s_X \mid S) = \sum_{k=s_X}^{\min\{n_X,s\}} \binom{n_X}{k}\binom{n_Y}{s-k} \Big/ \binom{n_X+n_Y}{s}$$

Therefore,

$$p(s_X) = P(S \geq s_X \mid s_X + s_Y) = \sum_{k=s_X}^{\min\{n_X,s_X+s_Y\}} \binom{n_X}{k}\binom{n_Y}{s_X+s_Y-k} \Big/ \binom{n_X+n_Y}{s_X+s_Y}$$

This defines the **Fisher's exact test**.

```
> fisher.test(...)
```

Example

	Coin 1	Coin 2
Head	520	680
Tail	480	520

```
> x <- matrix(c(520, 480, 680, 520), nr=2)
> fisher.test(x, alternative="two.sided")

data:  x
p-value = 0.03155
alternative hypothesis: true odds ratio is not equal to 1
...

> fisher.test(x, alternative="less")
data:  x
p-value = 0.01595
alternative hypothesis: true odds ratio is less than 1
...
```

Examples

	实验组	对照组
恶化的患者数量	0	24
没有恶化的患者数量	193	172

```
> x <- matrix(c(0, 193, 24, 172), nr=2);
> fisher.test(x, alternative="l");
```

$p = 3.407 \times 10^{-8}$

	实验组	对照组
发作冠心病的患者数量	12	20
不发作冠心病的患者数量	181	176

```
> x <- matrix(c(12, 181, 20, 176), nr=2);
> fisher.test(x, alternative="l");
```

$p = 0.1060$

Normal approximation

- Let

$$\hat{p}_x = \frac{s_x}{n_x}, \hat{p}_x = \frac{s_y}{n_y}, \hat{p} = \frac{s_x + s_y}{n_x + n_y},$$

- By the central limit theorem (CLT)

$$\frac{\hat{p}_x - p_x}{\sqrt{\hat{p}_x(1-\hat{p}_x)/n_x}} \to N(0,1), \frac{\hat{p}_y - p_y}{\sqrt{\hat{p}_y(1-\hat{p}_y)/n_y}} \to N(0,1), n \to +\infty$$

- Therefore ,with the assumption that p_x and p_y are independent, and under the null hypothesis $p_x=p_y=p$,

$$Z_n = \frac{\hat{p}_x - \hat{p}_y}{\sqrt{\hat{p}(1-\hat{p})(1/n_x + 1/n_y)}} \to N(0,1)$$

- This suggests the use of Zn as the test statistic and rejects H0 if

$$Z_n > z_\alpha$$

χ^2 approximation

- Since the square of a standard normal random variable is a χ^2 random variable. We have that

$$Z_n^2 = \frac{(\hat{p}_x - \hat{p}_y)^2}{\hat{p}(1-\hat{p})(1/n_x + 1/n_y)} \sim \chi_1^2, \quad n \to +\infty$$

- This suggests the use of Z_n^2 as the test statistic and reject H_0 if and only if

$$Z_n^2 > \chi_{1,\alpha}^2$$

Example

	Coin 1	Coin 2
Head	520	680
Tail	480	520

```
> x <- matrix(c(520, 480, 680, 520), nr=2)
> prop.test(x, alternative="two.sided")

2-sample test for equality of proportions with continuity correction

data:  x
X-squared = 4.6047, df = 1, p-value = 0.03188
alternative hypothesis: two.sided
...

> prop.test(x, alternative="less")
2-sample test for equality of proportions with continuity correction

data:  x
X-squared = 4.6047, df = 1, p-value = 0.01594
alternative hypothesis: less
...
```

Examples

	实验组	对照组
恶化的患者数量	0	24
没有恶化的患者数量	193	172

```
> x <- matrix(c(0, 193, 24, 172), nr=2);
> prop.test(x, alternative="l");
```

$$p = 3.407 \times 10^{-8}$$

	实验组	对照组
发作冠心病的患者数量	12	20
不发作冠心病的患者数量	181	176

```
> x <- matrix(c(12, 181, 20, 176), nr=2);
> fisher.test(x, alternative="l");
```

$$p = 0.1060$$

Expected frequency table

	Population X	Population Y	Total
Successes	s_x	S_y	$s = s_x + s_y$
Failures	f_x	f_y	$f = f_x + f_y$
Total	n_x	n_y	$n = n_x + n_y$

	Population X	Population Y	Total
Successes	$n_X\ s/n$	$n_Y\ s/n$	$s = s_x + s_y$
Failures	$n_X\ f/n$	$n_Y\ f/n$	$f = f_x + f_y$
Total	n_x	n_y	$n = n_x + n_y$

Difference in observed and expected

- Define the statistic

$$
\begin{aligned}
T &= \sum\sum \frac{\text{observed} - \text{expected}}{\text{expected}} \\
&= \frac{[s_x - n_x s/(n_x + n_y)]^2}{n_x s/(n_x + n_y)} + \frac{[s_y - n_y s/(n_x + n_y)]^2}{n_y s/(n_x + n_y)} \\
&+ \frac{[f_x - n_x f/(n_x + n_y)]^2}{n_x f/(n_x + n_y)} + \frac{[f_y - n_y f/(n_x + n_y)]^2}{n_y f/(n_x + n_y)} \\
&= \frac{(n_x\hat{p}_x - n_x\hat{p})^2}{n_x\hat{p}} + \frac{[n_x(1 - \hat{p}_x) - n_x(1 - p)]^2}{n_x(1 - \hat{p})} \\
&+ \frac{(n_y\hat{p}_y - n_y\hat{p})^2}{n_y\hat{p}} + \frac{[n_y(1 - \hat{p}_y) - n_y(1 - p)]^2}{n_y(1 - \hat{p})} \\
&= \frac{(n_x(\hat{p}_x - \hat{p})^2}{\hat{p}(1 - \hat{p})} + \frac{n_y(\hat{p}_y - \hat{p})^2}{\hat{p}(1 - \hat{p})}
\end{aligned}
$$

Difference in observed and expected

- Since

$$\hat{p} = \frac{s_x + s_y}{n_x + n_y} = \frac{n_x \hat{p}_x + n_y \hat{p}_y}{n_x + n_y}$$
$$T = \frac{n_x(\hat{p}_x - p)^2 + n_y(\hat{p}_y - p)^2}{\hat{p}(1-\hat{p})} = \frac{(\hat{p}_x - \hat{p}_y)^2)}{\hat{p}(1-\hat{p})(1/n_x + 1/n_y)}$$

- We have seen that

$$T = \frac{(\hat{p}_x - \hat{p}_y)^2}{\hat{p}(1-\hat{p})(\frac{1}{n_x} + \frac{1}{n_y})} \to \chi_1^2 \quad n \to +\infty$$

- This suggests the use of T as the test statistic and rejects if

$$T > \chi^2_{1,\alpha}$$

Test for two proportions

	实验组	对照组
恶化的患者的数量	0	24
没有恶化的患者的数量	193	172

```
> x <- matrix(c(0, 193, 24, 172), nr=2);
> chisq.test(x);
```

$p = 1.525 \times 10^{-6}$

	实验组	对照组
发作冠心病的患者数量	12	20
不发作冠心病的患者数量	181	176

```
> x <- matrix(c(12, 181, 20, 176), nr=2);
> chisq.test(x);
```

$p = 0.2127$

Test for two proportions

	实验组	对照组
恶化的患者数量	0	24
没有恶化的患者数量	193	172

```
> x <- matrix(c(0, 193, 24, 172), nr=2);
> prop.test(x, alternative="l");
```

$p = 7.627 \times 10^{-7}$

	实验组	对照组
发作冠心病的患者数量	12	20
不发作冠心病的患者数量	181	176

```
> x <- matrix(c(12, 181, 20, 176), nr=2);
> prop.test(x, alternative="l");
```

$p = 0.1063$

Tests of multiple proportions

Contingency table

	P_1	P_2		P_c	Total
C_1	n_{11}	n_{12}		n_{1c}	$n_{1\cdot}$
C_2	n_{21}	n_{22}		n_{2c}	$n_{2\cdot}$
			$\dots$		
C_r	n_{r1}	n_{r2}		n_{rc}	$n_{c\cdot}$
Total	$n_{\cdot 1}$	$n_{\cdot 2}$		$n_{\cdot c}$	$n_{\cdot\cdot}$

H_0 : The distributions over the categories are identical for all populations

versus

H_1 : The distributions over the categories are different for some populations

Expected numbers

Expected numbers

	P_1	P_2		P_c	Total
C_1	$n_{.1} \times n_{1.}/n_{..}$	$n_{.2} \times n_{1.}/n_{..}$		$n_{.c} \times n_{1.}/n_{..}$	$n_{1.}$
C_2	$n_{.1} \times n_{2.}/n_{..}$	$n_{.2} \times n_{2.}/n_{..}$		$n_{.c} \times n_{2.}/n_{..}$	$n_{2.}$
			$\cdots$		
C_r	$n_{.1} \times n_{r.}/n_{..}$	$n_{.2} \times n_{r.}/n_{..}$		$n_{.c} \times n_{r.}/n_{..}$	$n_{r.}$
Total	$n_{.1}$	$n_{.2}$		$n_{.c}$	$n_{..}$

$$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

Test

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \sim \chi^2_{(r-1)(c-1)}$$

A level α test would be:

$$\text{Reject } H_0 \text{ if and only if } \chi^2 > \chi^2_{(r-1)(c-1),\alpha}$$

The p-value is then

$$p = P(\chi^2_{(r-1)(c-1)} \geq \chi^2)$$

chisq.test

Another interpretation

Contingency table

	C_1	C_2		C_c	Total
R_1	n_{11}	n_{12}		n_{1c}	$n_{1\cdot}$
R_2	n_{21}	n_{22}		n_{2c}	$n_{2\cdot}$
			$\cdots$		
R_r	n_{r1}	n_{r2}		n_{rc}	$n_{c\cdot}$
Total	$n_{\cdot 1}$	$n_{\cdot 2}$		$n_{\cdot c}$	$n_{\cdot\cdot}$

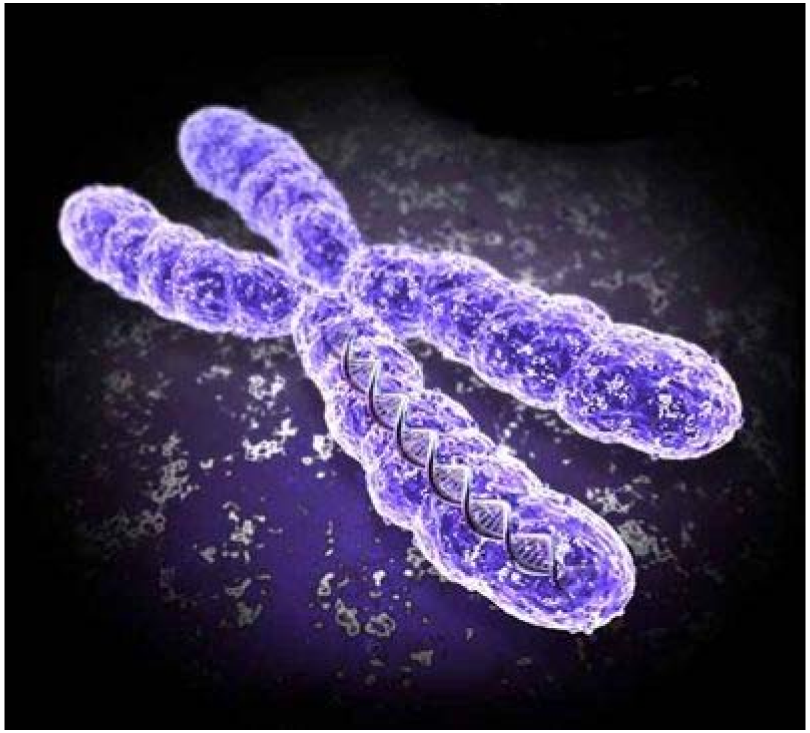
H_0 : The row parameter and the column parameter are independent

versus

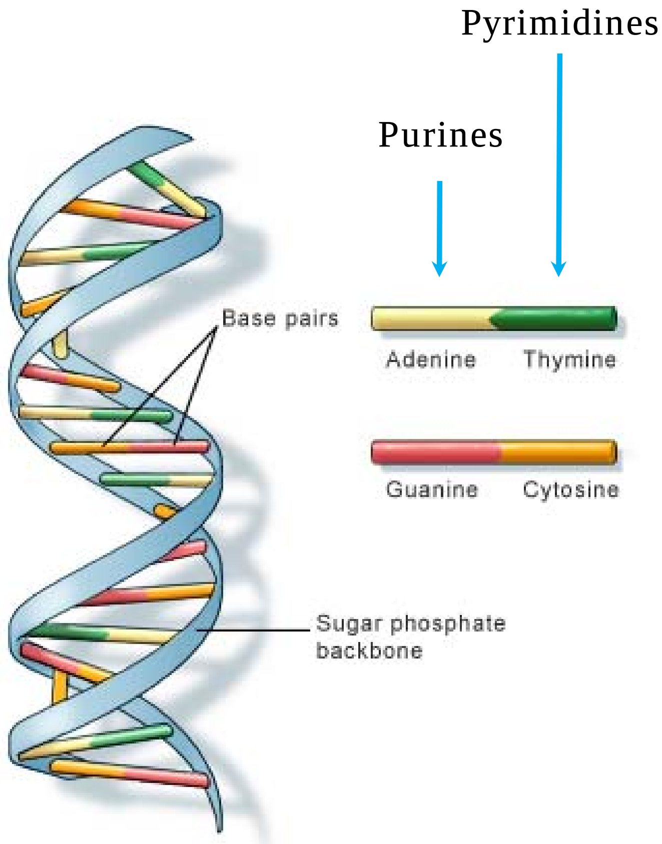
H_1 : The row parameter and the column parameter are NOT independent

	H_0	H_1	Exact test	Approximation
One-sample		$p \neq p_0$	Binomial exact test	Normal approximation χ^2 approximation
	$p = p_0$	$p > p_0$		
		$p < p_0$		
	$p \leq p_0$	$p > p_0$		
	$p \geq p_0$	$p < p_0$		
Two-sample		$p_X \neq p_Y$	Fisher exact test	Normal approximation χ^2 approximation
	$p_X = p_Y$	$p_X > p_Y$		
		$p_X < p_Y$		
	$p_X \leq p_Y$	$p_X > p_Y$		
	$p_X \geq p_Y$	$p_X < p_Y$		
Multi-sample	Identical distributions			
	Independence			

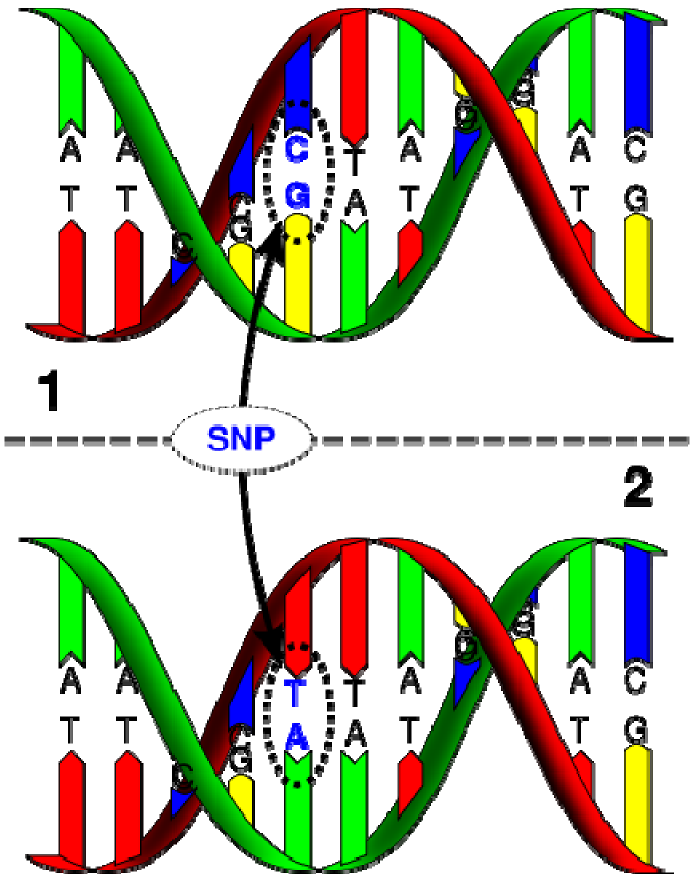
DNA



3 billion base-paires



Single nucleotide polymorphisms
SNP



Transition:

	A ↔ G:	Purine ↔ Purine
	C ↔ T:	Pyrimidine ↔ Pyrimidine

Transversion:

	A ↔ T:	Purine ↔ Pyrimidine
	A ↔ C:	Purine ↔ Pyrimidine
	G ↔ T:	Purine ↔ Pyrimidine
	G ↔ C:	Purine ↔ Pyrimidine

Theoretically,

FOUR possibilities for one chromosome

Reality,

TWO possibilities for one chromosome

THREE possibilities for two chromosomes

Each possibility is called a **genotype**

Case-control data

	Case 1		Case *N*	Control 1		Control *M*
SNP 1	0		1	0		1
SNP 2	1		1	0		0
SNP 3	1		1	0		0
	THREE possible genotypes					
SNP *L*	2		1	1		1

Whether a SNP is associated with the disease risk?

	Case 1		Case N	Control 1		Control M
SNP i	0		1	0		1

	Allele 0	Allele 1	Allele 2
Case	15	35	50
Control	75	15	10

Intuitive understanding via bar plot →

Quantitative measure ↓

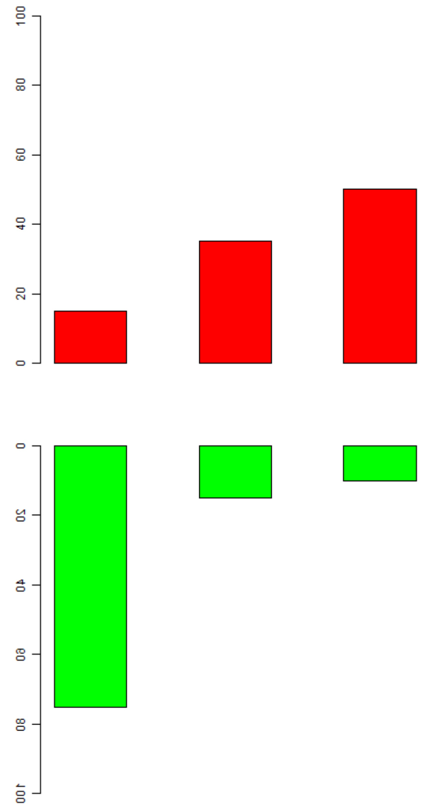
Fisher exact test

Chi-squared (χ^2) test

Cochran-Armitage test

Likelihood ratio test

…



Chi-squared test

Observed	Allele 1	Allele 2	Allele 3	Total
Case	o_{11}	o_{12}	o_{13}	n_1
Control	o_{21}	o_{22}	o_{23}	n_2
Total	o_1	o_2	o_3	n

Expected	Allele 1	Allele 2	Allele 3
Case	e_{11}	e_{12}	e_{13}
Control	e_{21}	e_{22}	e_{23}

$$e_{ij} = \frac{n_i o_j}{n}$$

$$\chi^2 = \sum_i \sum_j \frac{(o_{ij} - e_{ij})^2}{e_{ij}} \sim \chi^2_{(r-1)(c-1)} \sim \chi^2_2$$

Tests of Normal Mean

统计学方法及其应用 | 统计学基础 | 随机变量的函数

“A random variable is a quantity whose values are random and to which a probability distribution is assigned.”

Examples

	Active			Placebo		
	Pre	**Post**	**Post-pre**	**Pre**	**Post**	**Post-pre**
1	97	74	-23	94	79	-15
2	98	87	-11	103	101	-2
…						
192	90	78	-12	113	115	2
193	113	106	-7	112	109	-3
194				111	103	-8
195				101	88	-13
196				104	86	-18
	μ^{a}_{pre}	μ^{a}_{post}	$\mu^{a}_{\text{post-pre}}$	μ^{b}_{pre}	μ^{b}_{post}	$\mu^{b}_{\text{post-pre}}$

Hypotheses

- One-sample

$$H_0 : \mu^{\mathrm{a}}_{\mathrm{pre}} \leq 90 \quad \text{versus} \quad H_1 : \mu^{\mathrm{a}}_{\mathrm{pre}} > 90$$
$$H_0 : \mu^{\mathrm{a}}_{\mathrm{post}} = 90 \quad \text{versus} \quad H_1 : \mu^{\mathrm{a}}_{\mathrm{post}} > 90$$

- Paired-sample

$$H_0 : \mu^{\mathrm{a}}_{\mathrm{pre}} \leq \mu^{\mathrm{a}}_{\mathrm{post}} \quad \text{versus} \quad H_1 : \mu^{\mathrm{a}}_{\mathrm{pre}} > \mu^{\mathrm{a}}_{\mathrm{post}}$$
$$H_0 : \mu^{\mathrm{b}}_{\mathrm{pre}} = \mu^{\mathrm{b}}_{\mathrm{post}} \quad \text{versus} \quad H_1 : \mu^{\mathrm{b}}_{\mathrm{pre}} \neq \mu^{\mathrm{b}}_{\mathrm{post}}$$

- Two-sample

$$H_0 : \mu^{\mathrm{a}}_{\mathrm{pre}} = \mu^{\mathrm{b}}_{\mathrm{pre}} \quad \text{versus} \quad H_1 : \mu^{\mathrm{a}}_{\mathrm{pre}} \neq \mu^{\mathrm{b}}_{\mathrm{pre}}$$
$$H_0 : \mu^{\mathrm{a}}_{\mathrm{post}} = \mu^{\mathrm{b}}_{\mathrm{post}} \quad \text{versus} \quad H_1 : \mu^{\mathrm{a}}_{\mathrm{post}} < \mu^{\mathrm{b}}_{\mathrm{post}}$$

	H_0	H_1	σ^2 known	σ^2 unknown
One-sample		$\mu \neq \mu_0$	**One-sample z test**	**One-sample t test**
	$\mu = \mu_0$	$\mu > \mu_0$		
		$\mu < \mu_0$		
	$\mu \leq \mu_0$	$\mu > \mu_0$		
	$\mu \geq \mu_0$	$\mu < \mu_0$		
Two-sample		$\mu_X - \mu_Y \neq \delta_0$	**Two-sample z test**	**Two-sample t test**
	$\mu_X - \mu_Y = \delta_0$	$\mu_X - \mu_Y > \delta_0$		
		$\mu_X - \mu_Y < \delta_0$		
	$\mu_X - \mu_Y \leq \delta_0$	$\mu_X - \mu_Y > \delta_0$		
	$\mu_X - \mu_Y \geq \delta_0$	$\mu_X - \mu_Y < \delta_0$		
Paired-sample		$\mu_X - \mu_Y \neq \delta_0$	**Paired-sample z test**	**Paired-sample t test**
	$\mu_X - \mu_Y = \delta_0$	$\mu_X - \mu_Y > \delta_0$		
		$\mu_X - \mu_Y < \delta_0$		
	$\mu_X - \mu_Y \leq \delta_0$	$\mu_X - \mu_Y > \delta_0$		
	$\mu_X - \mu_Y \geq \delta_0$	$\mu_X - \mu_Y < \delta_0$		

One-sample tests of the normal mean

A random sample $X_1, \ldots X_n$ is observed from a **normal** population $N(\mu, \sigma^2)$, where σ^2 is **unknown**. We like to test

(1) $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$;

(2) $H_0 : \mu \leq \mu_0$ versus $H_1 : \mu > \mu_0$;

(3) $H_0 : \mu = \mu_0$ versus $H_1 : \mu > \mu_0$ (assume $\mu \geq \mu_0$);

(4) $H_0 : \mu \geq \mu_0$ versus $H_1 : \mu < \mu_0$;

(5) $H_0 : \mu = \mu_0$ versus $H_1 : \mu < \mu_0$ (assume $\mu \leq \mu_0$).

One-sample *t* test, two-sided

Hypothesis:

$$H_0 : \mu = \mu_0 \quad \text{versus} \quad H_1 : \mu \neq \mu_0$$

Test statistic:

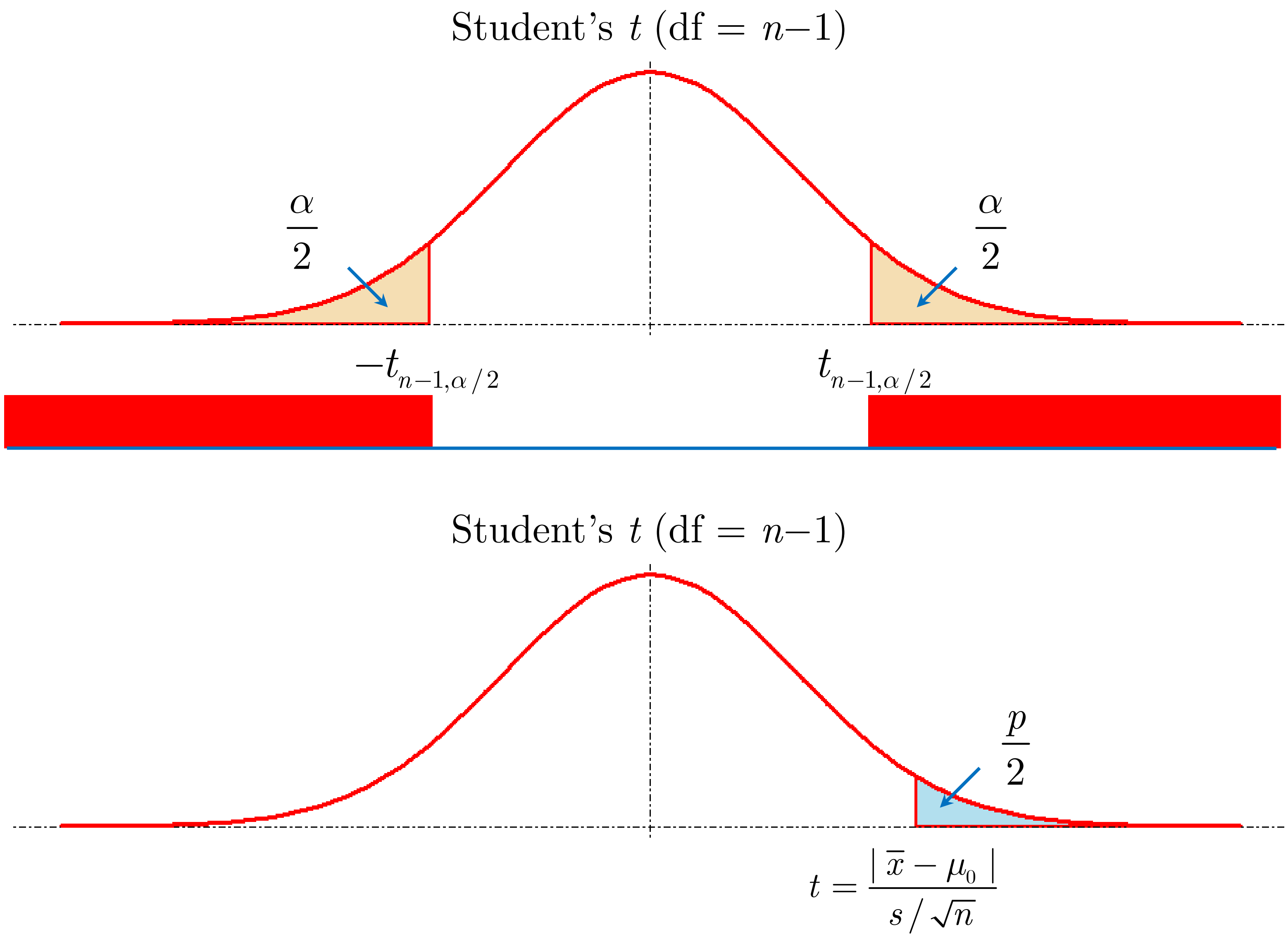
$$T = \frac{|\, \bar{X} - \mu_0 \,|}{S \,/\, \sqrt{n}}$$

Rejection region (level α):

$$R = \left\{ \mathbf{x} : \frac{|\, \bar{x} - \mu_0 \,|}{s \,/\, \sqrt{n}} > t_{n-1,\alpha/2} \right\}$$

p-value:

$$p(\mathbf{x}) = 2P\left(T_{n-1} \geq \frac{|\, \bar{x} - \mu_0 \,|}{s \,/\, \sqrt{n}} \right)$$


Student's t (df = $n-1$)
$\frac{\alpha}{2}$
$\frac{\alpha}{2}$
$-t_{n-1,\alpha/2}$
$t_{n-1,\alpha/2}$
Student's t (df = $n-1$)
$\frac{p}{2}$
$t = \frac{|\bar{x} - \mu_0|}{s/\sqrt{n}}$

One-sample *t* test, greater

Hypothesis:

$$H_0 : \mu \leq \mu_0 \quad \text{versus} \quad H_1 : \mu > \mu_0$$

$$H_0 : \mu = \mu_0 \quad \text{versus} \quad H_1 : \mu > \mu_0$$

Test statistic:

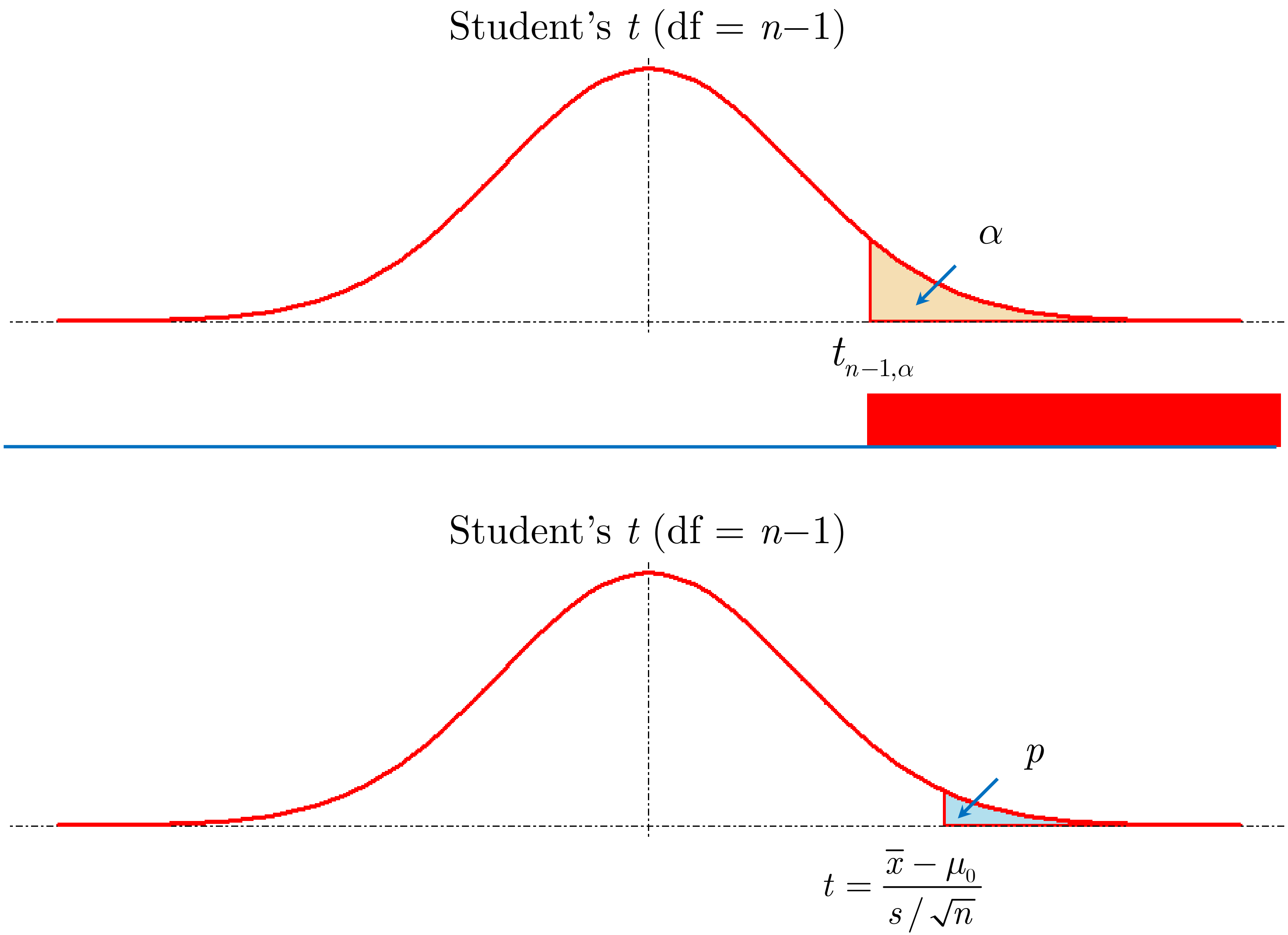
$$T = \frac{\bar{X} - \mu_0}{S \,/\, \sqrt{n}}$$

Rejection region (level α):

$$R = \left\{ \mathbf{x} : \frac{\bar{x} - \mu_0}{s \,/\, \sqrt{n}} > t_{n-1,\alpha} \right\}$$

p-value:

$$p(\mathbf{x}) = P\left(T_{n-1} \geq \frac{\bar{x} - \mu_0}{s \,/\, \sqrt{n}} \right)$$


Student's t (df = $n-1$)
α
$t_{n-1,\alpha}$
Student's t (df = $n-1$)
p
$t = \dfrac{\overline{x} - \mu_0}{s/\sqrt{n}}$

One-sample *t* test, less

Hypothesis:

$$H_0 : \mu \geq \mu_0 \quad \text{versus} \quad H_1 : \mu < \mu_0$$
$$H_0 : \mu = \mu_0 \quad \text{versus} \quad H_1 : \mu < \mu_0$$

Test statistic:

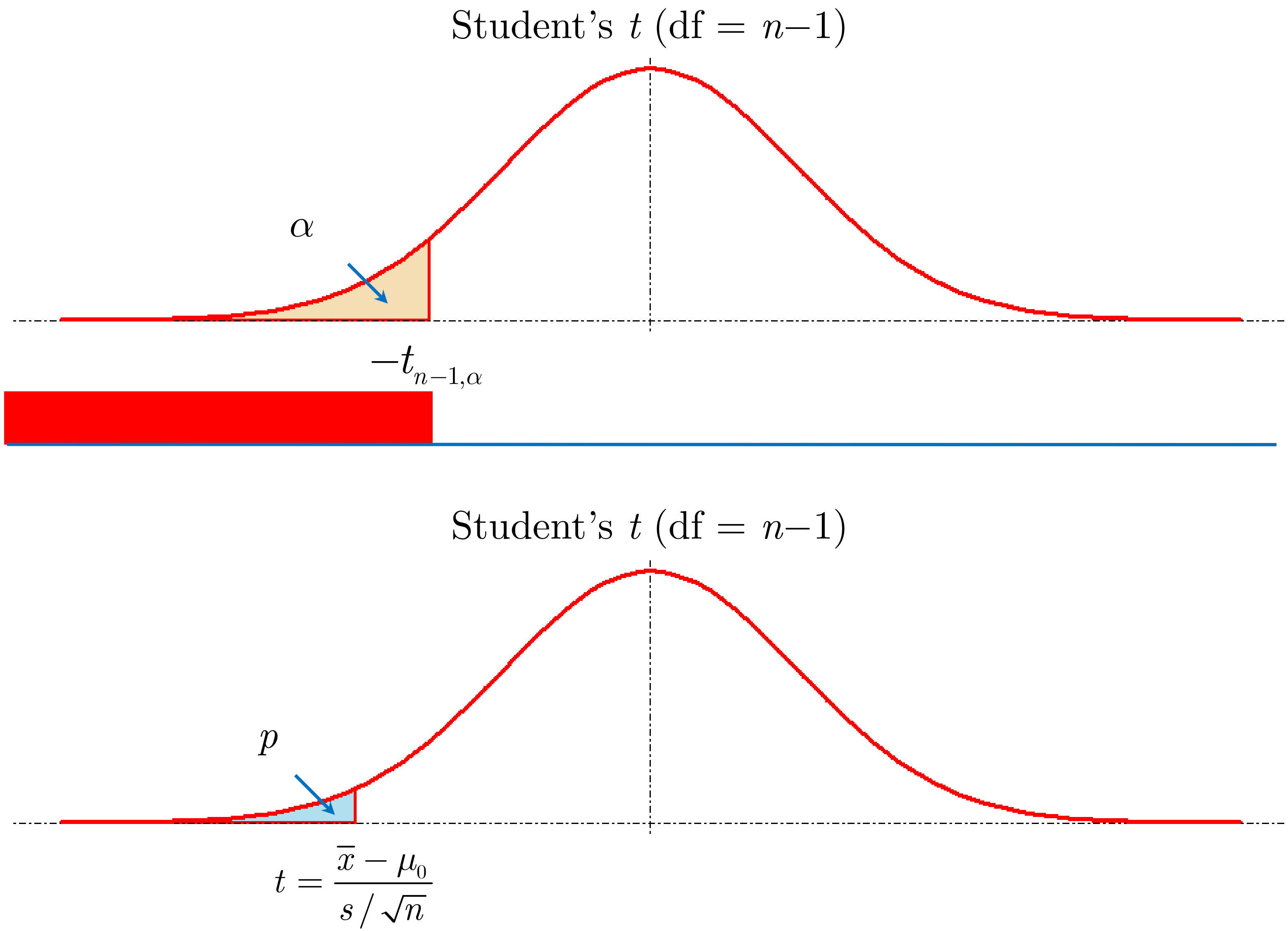
$$T = \frac{\bar{X} - \mu_0}{S \,/\, \sqrt{n}}$$

Rejection region (level α):

$$R = \left\{ \mathbf{x} : \frac{\bar{x} - \mu_0}{s \,/\, \sqrt{n}} < -t_{n-1,\alpha} \right\}$$

p-value:

$$p(\mathbf{x}) = P\left(T_{n-1} \leq \frac{\bar{x} - \mu_0}{s \,/\, \sqrt{n}} \right)$$



Student's t (df = $n-1$)
α
$-t_{n-1,\alpha}$
Student's t (df = $n-1$)
p
$t = \dfrac{\overline{x} - \mu_0}{s/\sqrt{n}}$

DBP of the active group (Begin)

100	99	107	96	104	96	94	105	97	91
105	114	93	98	102	98	104	102	100	93
113	99	99	97	102	102	105	114	110	109
115	91	94	112	95	102	95	91	104	111
101	107	114	101	102	99	108	91	94	110
106	114	93	95	95	90	94	96	94	112
110	96	114	97	99	115	106	103	103	106
94	93	92	90	104	113	102	93	95	92
114	105	97	93	95	95	102	115	104	104
108	90	94	109	109	95	95	105	111	109
90	114	94	96	98	105	114	100	113	115
115	92	99	103	93	99	102	114	102	96
98	109	96	112	115	98	109	96	105	106
92	93	93	91	100	114	106	115	96	95
113	99	110	110	104	114	102	92	92	95
108	110	101	99	113	111	111	103	100	91
96	94	91	108	102	93	90	93	109	108
108	114	111	90	104	100	90	95	109	101
98	113	103	96	110	96	97	96	94	91
97	113	110							

DBP: Diastolic Blood Pressure

DBP of the active group (End)

77	88	100	83	108	64	74	87	104	59
91	96	99	108	98	83	83	92	110	73
98	89	88	87	102	89	91	108	119	90
105	72	88	117	78	96	80	97	96	92
85	109	108	87	80	88	85	85	82	98
106	114	90	81	61	84	64	79	77	111
80	65	97	89	107	92	94	112	70	110
76	77	64	80	96	93	94	73	87	111
102	94	78	81	85	75	91	113	87	80
90	77	119	113	96	94	88	103	89	84
71	106	77	86	71	104	108	81	117	99
86	96	91	87	77	100	95	83	93	96
94	93	91	106	101	95	109	127	95	83
78	69	82	89	120	109	85	114	94	92
92	94	100	98	104	109	112	66	83	85
124	95	75	100	95	92	95	113	100	95
87	90	92	101	97	82	82	88	90	103
106	114	93	57	76	77	72	82	101	98
96	114	109	76	99	90	95	64	67	69
80	101	103							

DBP: Diastolic Blood Pressure

Practice

Test whether the patients really have hypertension

```
> t.test(active.begin, mu=90, alternative="g”)

        One Sample t-test

data:  active.begin
t = 22.1619, df = 192, p-value < 2.2e-16
alternative hypothesis: true mean is greater than 90
...

> t.test(active.end, mu=90, alternative="g”)

        One Sample t-test

data:  active.end
t = 2.0175, df = 192, p-value = 0.02252
alternative hypothesis: true mean is greater than 90
...
```

Paired-sample tests of the normal mean

A paired random sample $(X_1, Y_1), \ldots, (X_n, Y_n)$ is observed from a **bivariate normal** population $N(\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2, \rho)$.
We like to test

$$
\begin{array}{llll}
(1) & H_0 : \mu_X - \mu_Y = \delta_0 & \text{versus} & H_1 : \mu_X - \mu_Y \neq \delta_0; \\
(2) & H_0 : \mu_X - \mu_Y \leq \delta_0 & \text{versus} & H_1 : \mu_X - \mu_Y > \delta_0; \\
(3) & H_0 : \mu_X - \mu_Y = \delta_0 & \text{versus} & H_1 : \mu_X - \mu_Y > \delta_0; \\
(4) & H_0 : \mu_X - \mu_Y \geq \delta_0 & \text{versus} & H_1 : \mu_X - \mu_Y < \delta_0; \\
(5) & H_0 : \mu_X - \mu_Y = \delta_0 & \text{versus} & H_1 : \mu_X - \mu_Y < \delta_0.
\end{array}
$$

Paired-sample *t* test, greater

Hypothesis:

$$H_0 : \mu_X - \mu_Y \le \delta_0 \quad \text{versus} \quad H_1 : \mu_X - \mu_Y > \delta_0$$
$$H_0 : \mu_X - \mu_Y = \delta_0 \quad \text{versus} \quad H_1 : \mu_X - \mu_Y > \delta_0$$

Test statistic:

$$T = \frac{\bar{X} - \bar{Y} - \delta_0}{S_{X-Y} / \sqrt{n}}$$

Rejection region:

(level α) $$R = \left\{ \mathbf{x} : \frac{\bar{x} - \bar{y} - \delta_0}{s_{X-Y} / \sqrt{n}} > t_{n-1,\alpha} \right\}$$

p-value:

$$p(\mathbf{x}) = P\left(T_{n-1} \ge \frac{\bar{x} - \bar{y} - \delta_0}{s_{X-Y} / \sqrt{n}} \right)$$

$$S_{X-Y}^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left[(X_i - Y_i) - (\bar{X} - \bar{Y}) \right]^2$$

Paired-sample *t* test, less

Hypothesis:

$$H_0: \mu_X - \mu_Y \geq \delta_0 \quad \text{versus} \quad H_1: \mu_X - \mu_Y < \delta_0$$
$$H_0: \mu_X - \mu_Y = \delta_0 \quad \text{versus} \quad H_1: \mu_X - \mu_Y < \delta_0$$

Test statistic:

$$T = \frac{\bar{X} - \bar{Y} - \delta_0}{S_{X-Y} / \sqrt{n}}$$

Rejection region:

(level α) $$R = \left\{ \mathbf{x} : \frac{\bar{x} - \bar{y} - \delta_0}{s_{X-Y} / \sqrt{n}} < -t_{n-1,\alpha} \right\}$$

p-value:

$$p(\mathbf{x}) = P\left(T_{n-1} \leq \frac{\bar{x} - \bar{y} - \delta_0}{s_{X-Y} / \sqrt{n}} \right)$$

$$S_{X-Y}^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left[(X_i - Y_i) - (\bar{X} - \bar{Y}) \right]^2$$

Paired-sample *t* test, two-sided

Hypothesis:

$$H_0 : \mu_X - \mu_Y = \delta_0 \quad \text{versus} \quad H_1 : \mu_X - \mu_Y \neq \delta_0$$

Test statistic:

$$T = \frac{\left|\overline{X} - \overline{Y} - \delta_0\right|}{S_{X-Y} \,/\, \sqrt{n}}$$

Rejection region:

(level α) $$R = \left\{ \mathbf{x} : \frac{|\,\overline{x} - \overline{y} - \delta_0\,|}{s_{X-Y} \,/\, \sqrt{n}} > t_{n-1,\alpha/2} \right\}$$

p-value:

$$p(\mathbf{x}) = 2P\left(T_{n-1} \geq \frac{\left|\overline{x} - \overline{y} - \delta_0\right|}{s_{X-Y} \,/\, \sqrt{n}} \right)$$

$$S_{X-Y}^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left[(X_i - Y_i) - (\overline{X} - \overline{Y}) \right]^2$$

Active group *vs* Placebo group

1

98	113	103	96	110	96	97	96	94	91
97	113	110							

2

77	88	100	83	108	64	74	87	104	59
91	96	99	108	98	83	83	92	110	73
98	89	88	87	102	89	91	108	119	90
105	72	88	117	78	96	80	97	96	92
85	109	108	87	80	88	85	85	82	98
106	114	90	81	61	84	64	79	77	111
80	65	97	89	107	92	94	112	70	110
76	77	64	80	96	93	94	73	87	111
102	94	78	81	85	75	91	113	87	80
90	77	119	113	96	94	88	103	89	84
71	106	77	86	71	104	108	81	117	99
86	96	91	87	77	100	95	83	93	96
94	93	91	106	101	95	109	127	95	83
78	69	82	89	120	109	85	114	94	92
92	94	100	98	104	109	112	66	83	85
124	95	75	100	95	92	95	113	100	95
87	90	92	101	97	82	82	88	90	103
106	114	93	57	76	77	72	82	101	98
96	114	109	76	99	90	95	64	67	69
80	101	103							

3

109	97	112	115	113	110	105	114	115	90
92	104	109	104	115	90				

4

82	103	116	94	87	93	124	126	131	102
115	103	92	105	105	103	92	103	96	133
99	85	103	109	101	97	130	98	101	87
112	92	96	102	89	108	115	83	116	101
93	96	130	113	135	112	90	92	102	102
97	107	130	121	99	102	103	109	105	77
93	97	96	86	110	107	91	113	133	112
86	77	94	134	108	92	101	104	95	81
112	98	91	90	100	93	69	110	91	92
103	103	85	80	93	100	93	91	96	102
110	124	106	100	133	128	126	92	91	92
78	104	117	133	111	110	116	92	106	110
130	116	110	111	94	100	95	94	95	111
99	110	102	116	99	98	107	67	113	102
125	137	97	102	107	95	125	95	107	131
136	90	113	87	105	134	105	110	132	109
97	100	107	81	90	100	115	75	106	116
100	93	132	105	103	135	79	105	134	87
106	102	122	130	105	142	132	136	132	102
99	106	106	96	102	72				

Practice

Test whether the patients in the same group have lower blood pressure

```
> t.test(active.begin, active.end, "g", paired=T)

        Paired t-test

data:  active.begin and active.end
t = 12.3842, df = 192, p-value < 2.2e-16
alternative hypothesis: true difference in means is greater than 0
...

> t.test(placebo.begin, placebo.end, "t", paired=T)

         Paired t-test

data: placebo.begin and placebo.end
t = -1.0855, df = 195, p-value = 0.279
alternative hypothesis: true difference in means is not equal to 0
...
```

Two-sample tests of the normal mean

Two random samples $X_1,\dots X_m$ and $Y_1,\dots Y_n$ are obtained from **two normal** populations $N(\mu_X,\sigma_X^2)$ and $N(\mu_Y,\sigma_Y^2)$, respectively, where σ_X^2 and σ_Y^2 are **unknown** but an assumption that $\sigma_X^2=\sigma_Y^2$ holds. We like to test

(1) $H_0 : \mu_X - \mu_Y = \delta_0$ versus $H_1 : \mu_X - \mu_Y \neq \delta_0$;

(2) $H_0 : \mu_X - \mu_Y \leq \delta_0$ versus $H_1 : \mu_X - \mu_Y > \delta_0$;

(3) $H_0 : \mu_X - \mu_Y = \delta_0$ versus $H_1 : \mu_X - \mu_Y > \delta_0$;

(4) $H_0 : \mu_X - \mu_Y \geq \delta_0$ versus $H_1 : \mu_X - \mu_Y < \delta_0$;

(5) $H_0 : \mu_X - \mu_Y = \delta_0$ versus $H_1 : \mu_X - \mu_Y < \delta_0$.

Two-sample *t* test, greater

Hypothesis:

$$H_0 : \mu_X - \mu_Y \le \delta_0 \quad \text{versus} \quad H_1 : \mu_X - \mu_Y > \delta_0$$
$$H_0 : \mu_X - \mu_Y = \delta_0 \quad \text{versus} \quad H_1 : \mu_X - \mu_Y > \delta_0$$

Test statistic:

$$T = \frac{\overline{X} - \overline{Y} - \delta_0}{S_p \sqrt{1/m + 1/n}}$$

Regection region:

(level α)

$$R = \left\{ \mathbf{x} : \frac{\bar{x} - \bar{y} - \delta_0}{s_p \sqrt{1/m + 1/n}} > t_{m+n-2,\alpha} \right\}$$

p-value:

$$p(\mathbf{x}) = P\left(T_{m+n-2} \ge \frac{\bar{x} - \bar{y} - \delta_0}{s_p \sqrt{1/m + 1/n}} \right)$$

$$S_p^2 = \frac{(m-1)S_X^2 + (m-1)S_Y^2}{m+n-2}$$

Two-sample *t* test, less

Hypothesis:

$$H_0 : \mu_X - \mu_Y \geq \delta_0 \quad \text{versus} \quad H_1 : \mu_X - \mu_Y < \delta_0$$
$$H_0 : \mu_X - \mu_Y = \delta_0 \quad \text{versus} \quad H_1 : \mu_X - \mu_Y < \delta_0$$

Test statistic:

$$T = \frac{\overline{X} - \overline{Y} - \delta_0}{S_p \sqrt{1/m + 1/n}} \qquad S_p^2 = \frac{(m-1)S_X^2 + (m-1)S_Y^2}{m+n-2}$$

Regection region:

(level α)

$$R = \left\{ \mathbf{x} : \frac{\bar{x} - \bar{y} - \delta_0}{s_p \sqrt{1/m + 1/n}} < -t_{m+n-2,\alpha} \right\}$$

p-value:

$$p(\mathbf{x}) = P\left(T_{m+n-2} \leq \frac{\bar{x} - \bar{y} - \delta_0}{s_p \sqrt{1/m + 1/n}} \right)$$

Two-sample *t* test, two-sided

Hypothesis:

$$H_0 : \mu_X - \mu_Y = \delta_0 \quad \text{versus} \quad H_1 : \mu_X - \mu_Y \neq \delta_0$$

Test statistic:

$$T = \frac{\left|\overline{X} - \overline{Y} - \delta_0\right|}{S_p \sqrt{1/m + 1/n}} \qquad S_p^2 = \frac{(m-1)S_X^2 + (m-1)S_Y^2}{m+n-2}$$

Regection region:

(level α)

$$R = \left\{ \mathbf{x} : \frac{|\,\bar{x} - \bar{y} - \delta_0\,|}{s_p \sqrt{1/m + 1/n}} > t_{m+n-2,\alpha/2} \right\}$$

p-value:

$$p(\mathbf{x}) = 2P\left(T_{m+n-2} \geq \frac{|\,\bar{x} - \bar{y} - \delta_0\,|}{s_p \sqrt{1/m + 1/n}} \right)$$

Two-sample tests of the normal mean

Two random samples $X_1,\ldots X_m$ and $Y_1,\ldots Y_n$ are obtained from **two normal** populations $N(\mu_X,\sigma_X^2)$ and $N(\mu_Y,\sigma_Y^2)$, respectively, where σ_X^2 and σ_Y^2 are **unknown and unequal**. We like to test

$$
\begin{array}{lllll}
(1) & H_0:\mu_X-\mu_Y = \delta_0 & \text{versus} & H_1:\mu_X-\mu_Y \neq \delta_0; \\
(2) & H_0:\mu_X-\mu_Y \leq \delta_0 & \text{versus} & H_1:\mu_X-\mu_Y > \delta_0; \\
(3) & H_0:\mu_X-\mu_Y = \delta_0 & \text{versus} & H_1:\mu_X-\mu_Y > \delta_0; \\
(4) & H_0:\mu_X-\mu_Y \geq \delta_0 & \text{versus} & H_1:\mu_X-\mu_Y < \delta_0; \\
(5) & H_0:\mu_X-\mu_Y = \delta_0 & \text{versus} & H_1:\mu_X-\mu_Y < \delta_0.
\end{array}
$$

Two-sample *t* test, greater (approximated)

Hypothesis:

$$H_0 : \mu_X \leq \mu_Y \quad \text{versus} \quad H_1 : \mu_X > \mu_Y$$
$$H_0 : \mu_X = \mu_Y \quad \text{versus} \quad H_1 : \mu_X > \mu_Y$$

Test statistic:

$$T = \frac{\overline{X} - \overline{Y} - \delta_0}{\sqrt{S_X^2 / m + S_Y^2 / n}},$$

Regection region:

(level α)

$$R = \left\{ \mathbf{x} : \frac{\bar{x} - \bar{y} - \delta_0}{\sqrt{s_X^2 / m + s_Y^2 / n}} > t_{k,\alpha} \right\}$$

p-value:

$$p(\mathbf{x}) = P\left(T_k \geq \frac{\bar{x} - \bar{y} - \delta_0}{\sqrt{s_X^2 / m + s_Y^2 / n}} \right)$$

$$K = \frac{\left(\frac{S_X^2}{m} + \frac{S_Y^2}{n} \right)^2}{\frac{S_X^4}{m^2(m-1)} + \frac{S_Y^4}{n^2(n-1)}}$$

Two-sample *t* test, less (approximated)

Hypothesis:

$$H_0 : \mu_X \geq \mu_Y \quad \text{versus} \quad H_1 : \mu_X < \mu_Y$$
$$H_0 : \mu_X = \mu_Y \quad \text{versus} \quad H_1 : \mu_X < \mu_Y$$

Test statistic:

$$T = \frac{\overline{X} - \overline{Y} - \delta_0}{\sqrt{S_X^2 / m + S_Y^2 / n}}, \qquad K = \frac{\left(\dfrac{S_X^2}{m} + \dfrac{S_Y^2}{n}\right)^2}{\dfrac{S_X^4}{m^2(m-1)} + \dfrac{S_Y^4}{n^2(n-1)}}$$

Regection region:

(level α)

$$R = \left\{ \mathbf{x} : \frac{\bar{x} - \bar{y} - \delta_0}{\sqrt{s_X^2 / m + s_Y^2 / n}} < -t_{k,\alpha} \right\}$$

p-value:

$$p(\mathbf{x}) = P\left(T_k \leq \frac{\bar{x} - \bar{y} - \delta_0}{\sqrt{s_X^2 / m + s_Y^2 / n}} \right)$$

Two-sample *t* test, two-sided (approximated)

Hypothesis:

$$H_0 : \mu_X - \mu_Y = \delta_0 \quad \text{versus} \quad H_1 : \mu_X - \mu_Y \neq \delta_0$$

Test statistic:

$$T = \frac{\left|\overline{X} - \overline{Y} - \delta_0\right|}{\sqrt{S_X^2 / m + S_Y^2 / n}}, \qquad K = \frac{\left(\frac{S_X^2}{m} + \frac{S_Y^2}{n}\right)^2}{\frac{S_X^4}{m^2(m-1)} + \frac{S_Y^4}{n^2(n-1)}}$$

Regection region:

(level α)

$$R = \left\{ \mathbf{x} : \frac{|\,\bar{x} - \bar{y} - \delta_0\,|}{\sqrt{s_X^2 / m + s_Y^2 / n}} > t_{k,\alpha/2} \right\}$$

p-value:

$$p(\mathbf{x}) = 2P\left(T_k \geq \frac{|\,\bar{x} - \bar{y} - \delta_0\,|}{\sqrt{s_X^2 / m + s_Y^2 / n}} \right)$$

DBP of the active group (Begin)

100	99	107	96	104	96	94	105	97	91
105	114	93	98	102	98	104	102	100	93
113	99	99	97	102	102	105	114	110	109
115	91	94	112	95	102	95	91	104	111
101	107	114	101	102	99	108	91	94	110
106	114	93	95	95	90	94	96	94	112
110	96	114	97	99	115	106	103	103	106
94	93	92	90	104	113	102	93	95	92
114	105	97	93	95	95	102	115	104	104
108	90	94	109	109	95	95	105	111	109
90	114	94	96	98	105	114	100	113	115
115	92	99	103	93	99	102	114	102	96
98	109	96	112	115	98	109	96	105	106
92	93	93	91	100	114	106	115	96	95
113	99	110	110	104	114	102	92	92	95
108	110	101	99	113	111	111	103	100	91
96	94	91	108	102	93	90	93	109	108
108	114	111	90	104	100	90	95	109	101
98	113	103	96	110	96	97	96	94	91
97	113	110							

DBP: Diastolic Blood Pressure

DBP of the placebo group (Begin)

97	105	110	103	90	94	115	111	114	99
105	113	110	103	90	106	93	93	91	113
113	91	100	99	104	96	114	98	101	92
106	106	95	94	98	98	109	93	112	104
105	91	113	111	115	109	98	108	114	115
103	102	113	113	104	110	112	97	112	98
103	99	100	104	104	115	99	103	113	107
97	96	107	115	114	102	103	96	93	94
101	90	91	107	100	109	92	90	112	98
99	108	97	97	113	106	91	96	91	100
110	109	105	96	115	113	107	109	96	102
92	96	113	113	112	100	104	97	101	115
110	109	103	115	94	102	94	94	94	111
99	110	112	109	95	98	107	93	111	96
105	114	99	91	111	102	105	91	104	111
113	92	102	91	112	114	101	107	112	94
95	110	105	97	91	106	112	94	99	110
93	91	110	101	109	115	114	108	111	94
109	97	112	115	113	110	105	114	115	90
92	104	109	104	115	90				

DBP of the active group (End)

77	88	100	83	108	64	74	87	104	59
91	96	99	108	98	83	83	92	110	73
98	89	88	87	102	89	91	108	119	90
105	72	88	117	78	96	80	97	96	92
85	109	108	87	80	88	85	85	82	98
106	114	90	81	61	84	64	79	77	111
80	65	97	89	107	92	94	112	70	110
76	77	64	80	96	93	94	73	87	111
102	94	78	81	85	75	91	113	87	80
90	77	119	113	96	94	88	103	89	84
71	106	77	86	71	104	108	81	117	99
86	96	91	87	77	100	95	83	93	96
94	93	91	106	101	95	109	127	95	83
78	69	82	89	120	109	85	114	94	92
92	94	100	98	104	109	112	66	83	85
124	95	75	100	95	92	95	113	100	95
87	90	92	101	97	82	82	88	90	103
106	114	93	57	76	77	72	82	101	98
96	114	109	76	99	90	95	64	67	69
80	101	103							

DBP of the placebo group (End)

82	103	116	94	87	93	124	126	131	102
115	103	92	105	105	103	92	103	96	133
99	85	103	109	101	97	130	98	101	87
112	92	96	102	89	108	115	83	116	101
93	96	130	113	135	112	90	92	102	102
97	107	130	121	99	102	103	109	105	77
93	97	96	86	110	107	91	113	133	112
86	77	94	134	108	92	101	104	95	81
112	98	91	90	100	93	69	110	91	92
103	103	85	80	93	100	93	91	96	102
110	124	106	100	133	128	126	92	91	92
78	104	117	133	111	110	116	92	106	110
130	116	110	111	94	100	95	94	95	111
99	110	102	116	99	98	107	67	113	102
125	137	97	102	107	95	125	95	107	131
136	90	113	87	105	134	105	110	132	109
97	100	107	81	90	100	115	75	106	116
100	93	132	105	103	135	79	105	134	87
106	102	122	130	105	142	132	136	132	102
99	106	106	96	102	72				

Practice

Test whether the two group of patients have similar hypertension strength

```
> t.test(active.begin, placebo.begin, "t", var.equal=T)

        Two Sample t-test

data:  active.begin and placebo.end
t = 1.1862, df = 387, p-value = 0.2363
alternative hypothesis: true difference in means is not equal to 0
...

> t.test(active.begin, placebo.begin, "l", var.equal=T)

        Two Sample t-test

data: active.begin and placebo.end
t = -7.5325, df = 387, p-value = 1.774e-13
alternative hypothesis: true difference in means is less than 0
...
```

Practice

Test whether the two group of patients have similar hypertension strength

```
> t.test(active.begin, placebo.begin, "t", var.equal=F)

         Welch Two Sample t-test

data:  active.begin and placebo.begin
t = 1.1862, df = 386.863, p-value = 0.2363
alternative hypothesis: true difference in means is not equal to 0
...

> t.test(active.end, placebo.end, "l", var.equal=F)

         Welch Two Sample t-test

data: active.end and placebo.end
t = -7.5291, df = 384.883, p-value = 1.833e-13
alternative hypothesis: true difference in means is less than 0
...
```

Tests of Normal Variance

统计学方法及其应用 | 统计学基础 | 随机变量的函数

"A random variable is a quantity whose values are random and to which a probability distribution is assigned."

	H_0	H_1	μ known	μ unknown
One-sample		$\sigma^2 \neq \sigma_0^2$	χ^2 test	χ^2 test
	$\sigma^2 = \sigma_0^2$	$\sigma^2 > \sigma_0^2$		
		$\sigma^2 < \sigma_0^2$		
	$\sigma^2 \leq \sigma_0^2$	$\sigma^2 > \sigma_0^2$		
	$\sigma^2 \geq \sigma_0^2$	$\sigma^2 < \sigma_0^2$		
Two-sample		$\sigma_X^2 / \sigma_Y^2 \neq \lambda_0$	F test	F test
	$\sigma_X^2 / \sigma_Y^2 = \lambda_0$	$\sigma_X^2 / \sigma_Y^2 > \lambda_0$		
		$\sigma_X^2 / \sigma_Y^2 < \lambda_0$		
	$\sigma_X^2 / \sigma_Y^2 \leq \lambda_0$	$\sigma_X^2 / \sigma_Y^2 > \lambda_0$		
	$\sigma_X^2 / \sigma_Y^2 \geq \lambda_0$	$\sigma_X^2 / \sigma_Y^2 < \lambda_0$		

One-sample test, greater

Hypothesis:

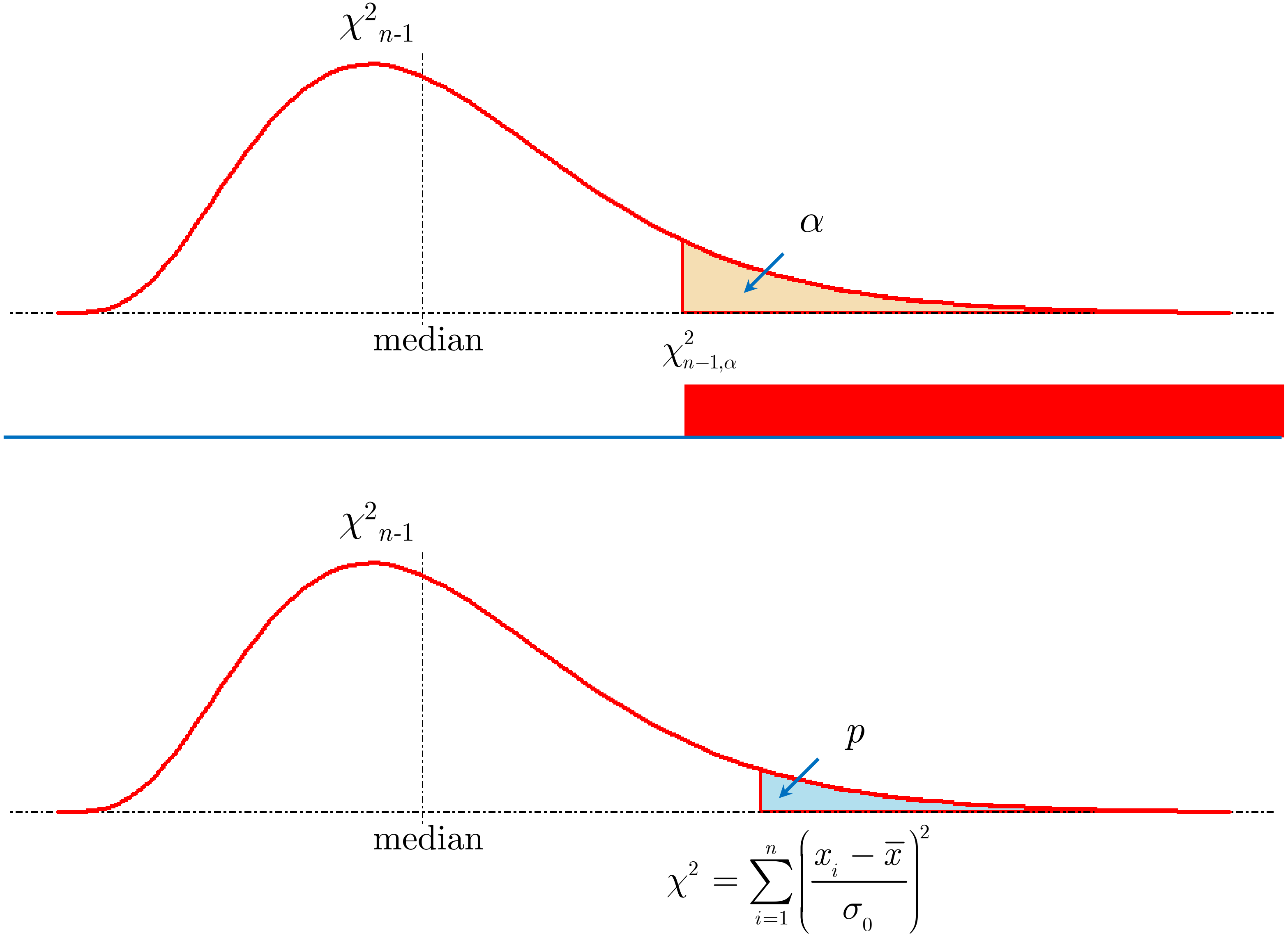
$$H_0 : \sigma^2 = \sigma_0^2 \quad \text{versus} \quad H_1 : \sigma^2 > \sigma_0^2$$

Test statistic:

$$\chi^2 = \sum_{i=1}^{n} \left(\frac{X_i - \bar{X}}{\sigma_0} \right)^2, \qquad \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$$

p-value:

$$p(\mathbf{x}) = P\left(\chi^2_{n-1} \geq \sum_{i=1}^{n} \left(\frac{x_i - \bar{x}}{\sigma_0} \right)^2 \right)$$



$\chi^2{}_{n\text{-}1}$
α
median
$\chi^2_{n-1,\alpha}$
$\chi^2{}_{n\text{-}1}$
p
median
$\chi^2 = \sum_{i=1}^{n}\left(\frac{x_i - \overline{x}}{\sigma_0}\right)^2$

One-sample test, less

Hypothesis:

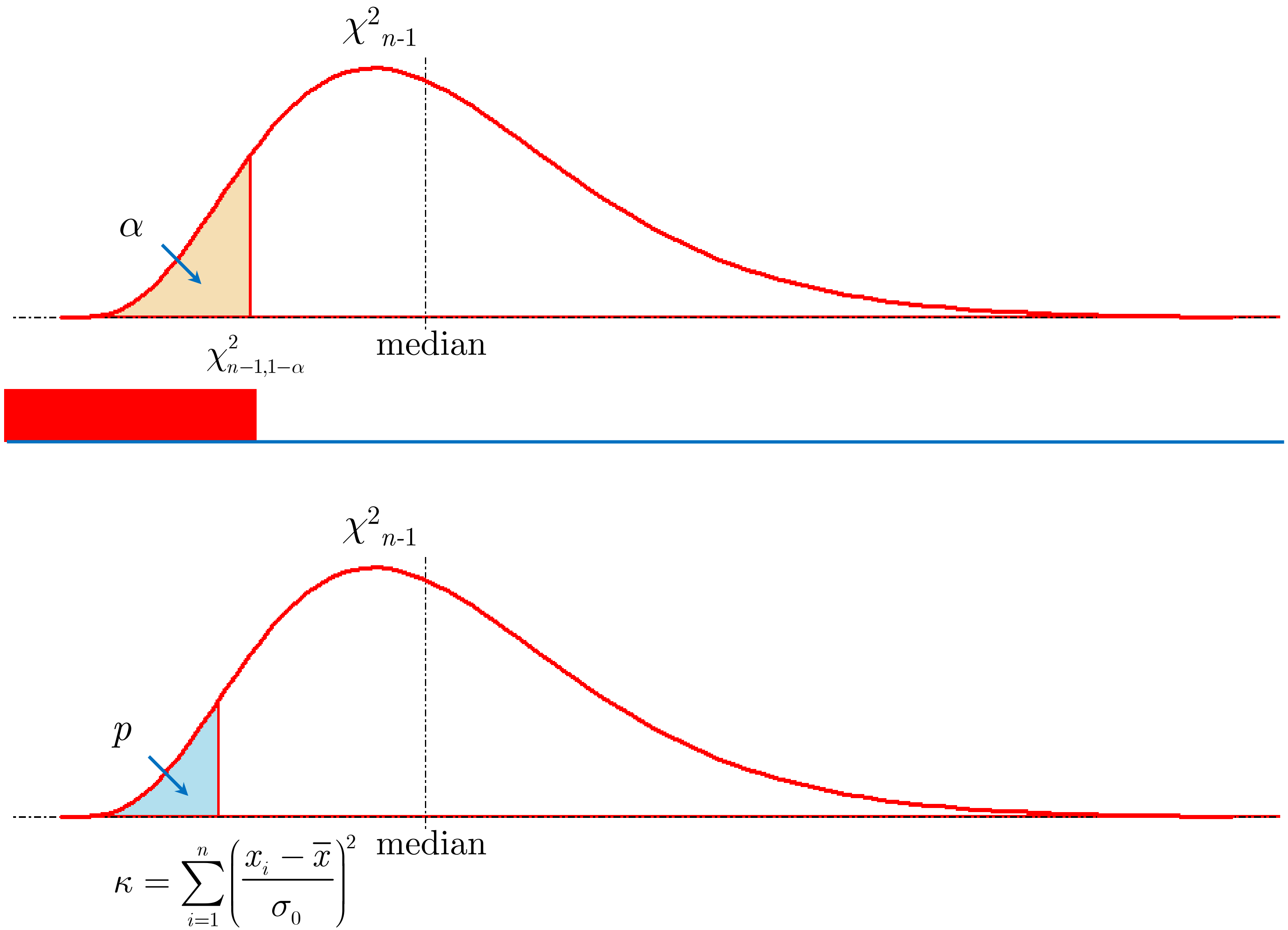
$$H_0 : \sigma^2 = \sigma_0^2 \quad \text{versus} \quad H_1 : \sigma^2 < \sigma_0^2$$

Test statistic:

$$\chi^2 = \sum_{i=1}^{n} \left(\frac{X_i - \bar{X}}{\sigma_0} \right)^2$$

p-value:

$$p(\mathbf{x}) = P\left(\chi^2_{n-1} \leq \sum_{i=1}^{n} \left(\frac{x_i - \bar{x}}{\sigma_0} \right)^2 \right)$$



$\chi^2_{n\text{-}1}$
α
median
$\chi^2_{n-1,1-\alpha}$
$\chi^2_{n\text{-}1}$
p
median
$\kappa = \sum_{i=1}^{n} \left(\frac{x_i - \overline{x}}{\sigma_0} \right)^2$

One-sample test, two-sided

Hypothesis:

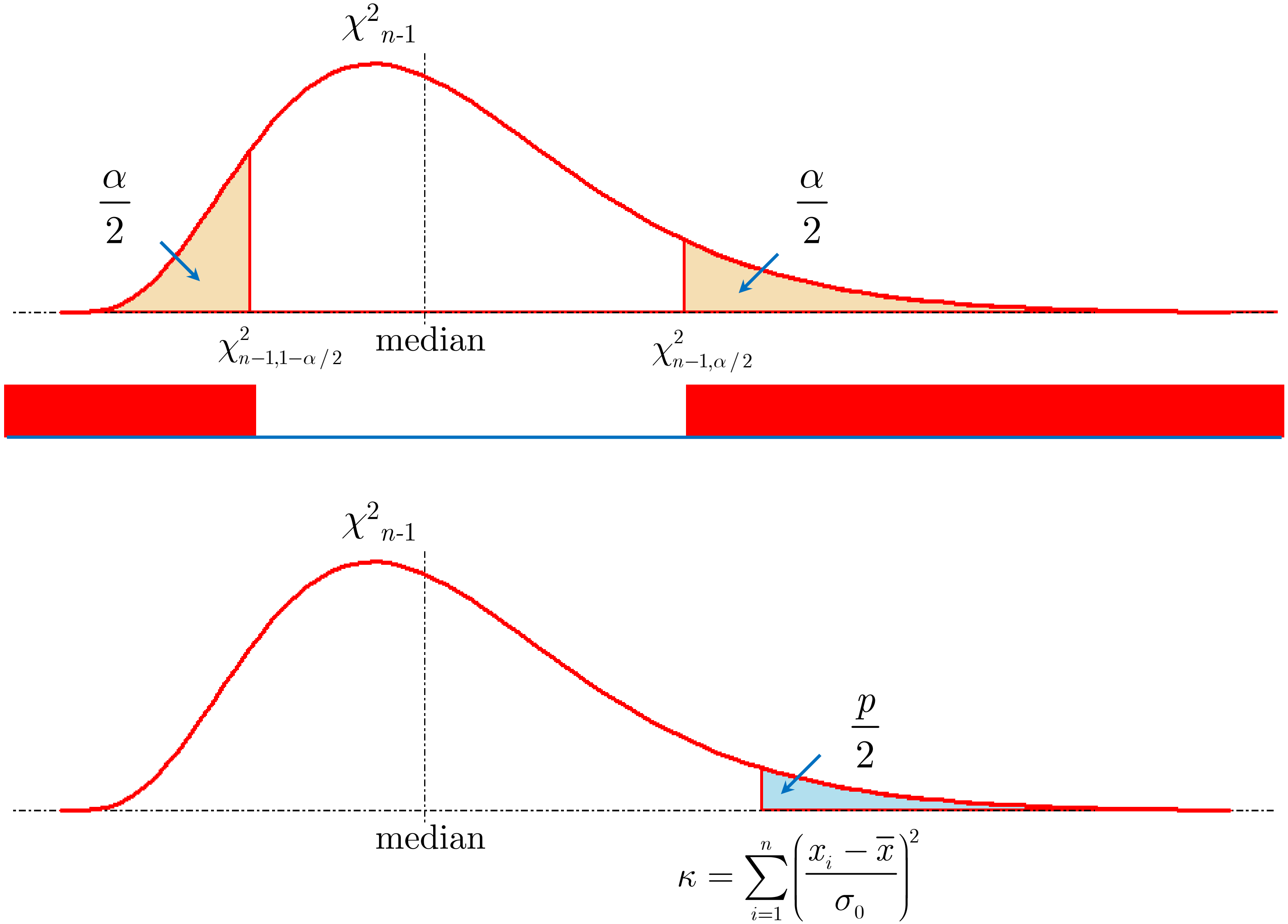
$$H_0 : \sigma^2 = \sigma_0^2 \quad \text{versus} \quad H_1 : \sigma^2 \neq \sigma_0^2$$

Test statistic:

$$K^2 = \sum_{i=1}^{n} \left(\frac{X_i - \overline{X}}{\sigma_0} \right)^2$$

p-value:

$$p(\mathbf{x}) = 2 \min \left\{ P\left(\chi_{n-1}^2 \geq \sum_{i=1}^{n} \left(\frac{x_i - \overline{x}}{\sigma_0} \right)^2 \right), P\left(\chi_{n-1}^2 \leq \sum_{i=1}^{n} \left(\frac{x_i - \overline{x}}{\sigma_0} \right)^2 \right) \right\}$$



$\chi^2{}_{n\text{-}1}$
$\frac{\alpha}{2}$
$\frac{\alpha}{2}$
$\chi^2_{n-1,1-\alpha/2}$
median
$\chi^2_{n-1,\alpha/2}$
$\chi^2{}_{n\text{-}1}$
$\frac{p}{2}$
median
$\kappa = \sum_{i=1}^{n}\left(\frac{x_i - \overline{x}}{\sigma_0}\right)^2$

Two-sample tests of the normal variances

Two random samples $X_1,\ldots,X_m$ and $Y_1,\ldots,Y_n$ are obtained from two **normal** population $N(\mu_X,\sigma_X^2)$ and $N(\mu_Y,\sigma_Y^2)$, respectively, where μ_X and μ_Y are **unknown**.

We like to test

(1) $H_0 : \sigma_X^2 / \sigma_Y^2 = \lambda_0$ versus $H_1 : \sigma_X^2 / \sigma_Y^2 > \lambda_0$;

(2) $H_0 : \sigma_X^2 / \sigma_Y^2 = \lambda_0$ versus $H_1 : \sigma_X^2 / \sigma_Y^2 < \lambda_0$;

(3) $H_0 : \sigma_X^2 / \sigma_Y^2 = \lambda_0$ versus $H_1 : \sigma_X^2 / \sigma_Y^2 \neq \lambda_0$.

Two-sample test, greater

Hypothesis:

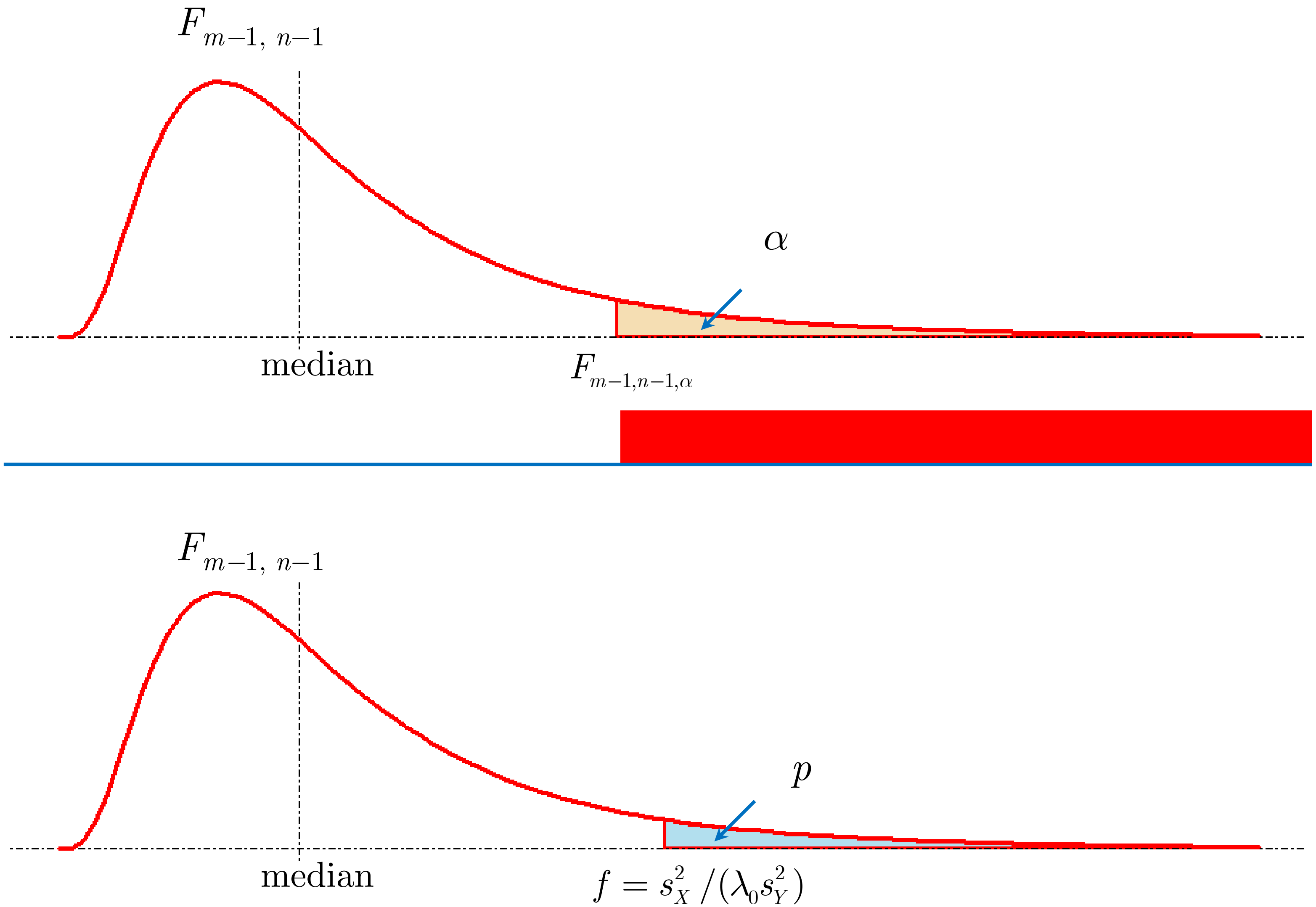
$$H_0 : \sigma_X^2 / \sigma_Y^2 = \lambda_0 \quad \text{versus} \quad H_1 : \sigma_X^2 / \sigma_Y^2 > \lambda_0$$

Test statistic:

$$F = \frac{1}{\lambda_0} \frac{S_X^2}{S_Y^2}, \qquad \frac{S_X^2 / S_Y^2}{\sigma_X^2 / \sigma_Y^2} \sim F_{m-1,n-1}$$

p-value:

$$p(\mathbf{x}) = P\left(F_{m-1,n-1} \geq \frac{1}{\lambda_0} \frac{s_X^2}{s_Y^2} \right)$$



$F_{m-1,\, n-1}$
α
median
$F_{m-1,n-1,\alpha}$
$F_{m-1,\, n-1}$
p
median
$f = s_X^2 / (\lambda_0 s_Y^2)$

Two-sample test, less

Hypothesis:

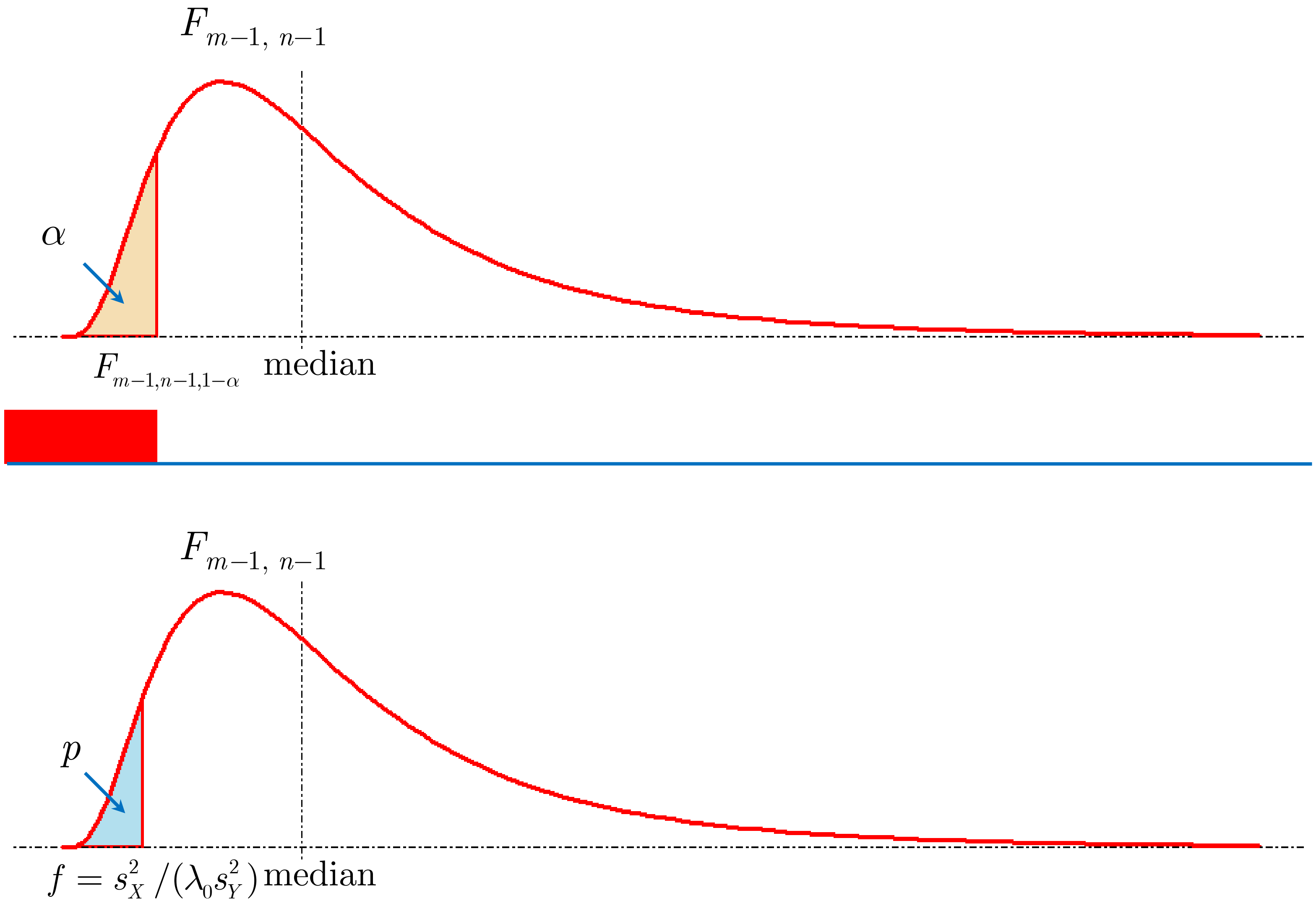
$$H_0 : \sigma_X^2 / \sigma_Y^2 = \lambda_0 \quad \text{versus} \quad H_1 : \sigma_X^2 / \sigma_Y^2 < \lambda_0$$

Test statistic:

$$F = \frac{1}{\lambda_0} \frac{S_X^2}{S_Y^2}$$

p-value:

$$p(\mathbf{x}) = P\left(F_{m-1,n-1} \leq \frac{1}{\lambda_0} \frac{s_X^2}{s_Y^2} \right)$$



$F_{m-1,\ n-1}$
α
$F_{m-1,n-1,1-\alpha}$
median
$F_{m-1,\ n-1}$
p
$f = s_X^2/(\lambda_0 s_Y^2)$
median

Two-sample test, two-sided

Hypothesis:

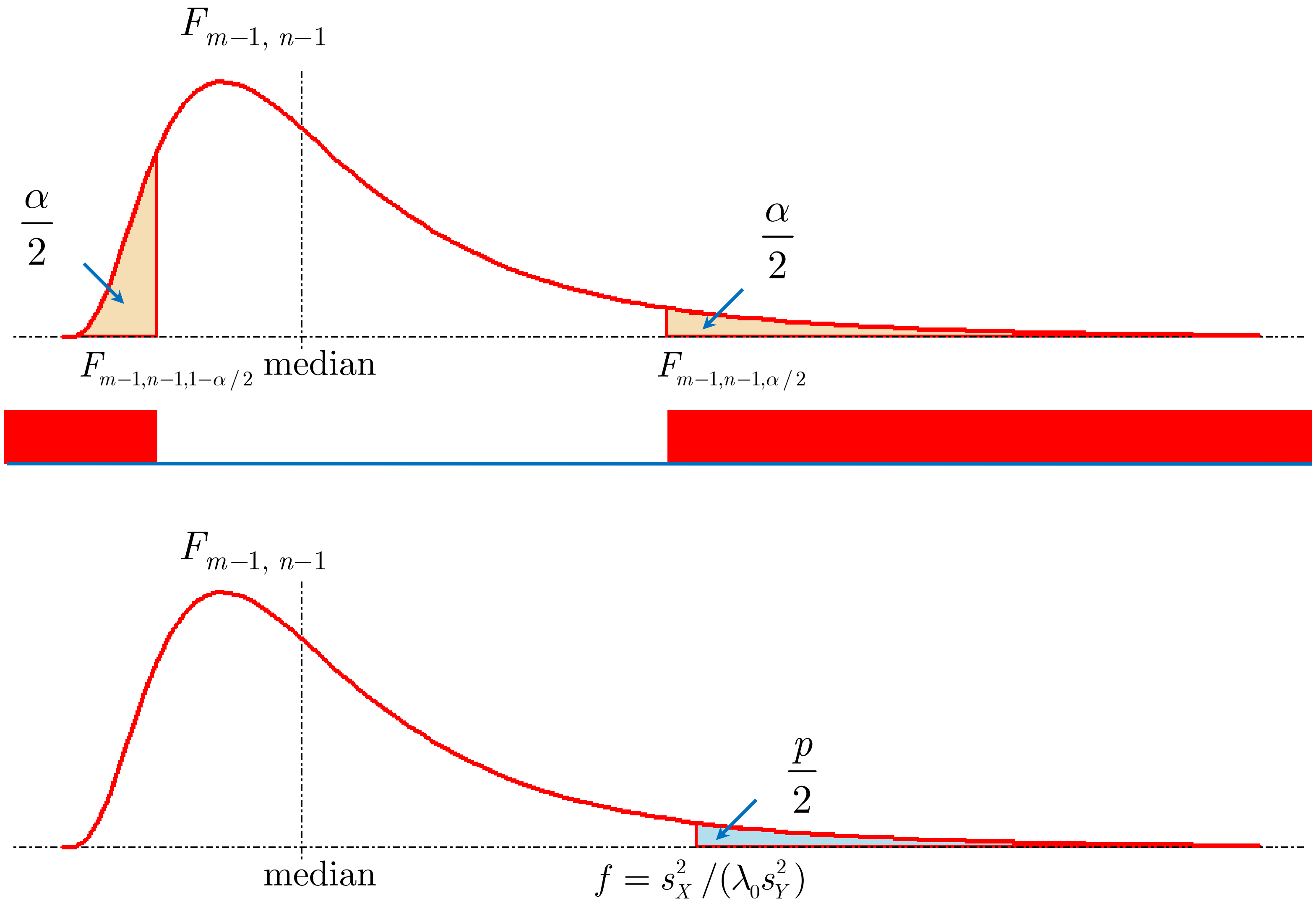
$$H_0 : \sigma_X^2 / \sigma_Y^2 = \lambda_0 \quad \text{versus} \quad H_1 : \sigma_X^2 / \sigma_Y^2 \neq \lambda_0$$

Test statistic:

$$F = \frac{1}{\lambda_0} \frac{S_X^2}{S_Y^2}$$

p-value:

$$p(\mathbf{x}) = 2 \min \left\{ P\left(F_{m-1,n-1} \geq \frac{1}{\lambda_0} \frac{s_X^2}{s_Y^2} \right), P\left(F_{m-1,n-1} \leq \frac{1}{\lambda_0} \frac{s_X^2}{s_Y^2} \right) \right\}$$



$F_{m-1,\ n-1}$
$\frac{\alpha}{2}$
$\frac{\alpha}{2}$
$F_{m-1,n-1,1-\alpha/2}$
median
$F_{m-1,n-1,\alpha/2}$
$F_{m-1,\ n-1}$
$\frac{p}{2}$
median
$f = s_X^2/(\lambda_0 s_Y^2)$

DBP of the active group (Begin)

100	99	107	96	104	96	94	105	97	91
105	114	93	98	102	98	104	102	100	93
113	99	99	97	102	102	105	114	110	109
115	91	94	112	95	102	95	91	104	111
101	107	114	101	102	99	108	91	94	110
106	114	93	95	95	90	94	96	94	112
110	96	114	97	99	115	106	103	103	106
94	93	92	90	104	113	102	93	95	92
114	105	97	93	95	95	102	115	104	104
108	90	94	109	109	95	95	105	111	109
90	114	94	96	98	105	114	100	113	115
115	92	99	103	93	99	102	114	102	96
98	109	96	112	115	98	109	96	105	106
92	93	93	91	100	114	106	115	96	95
113	99	110	110	104	114	102	92	92	95
108	110	101	99	113	111	111	103	100	91
96	94	91	108	102	93	90	93	109	108
108	114	111	90	104	100	90	95	109	101
98	113	103	96	110	96	97	96	94	91
97	113	110							

DBP: Diastolic Blood Pressure

DBP of the placebo group (Begin)

97	105	110	103	90	94	115	111	114	99
105	113	110	103	90	106	93	93	91	113
113	91	100	99	104	96	114	98	101	92
106	106	95	94	98	98	109	93	112	104
105	91	113	111	115	109	98	108	114	115
103	102	113	113	104	110	112	97	112	98
103	99	100	104	104	115	99	103	113	107
97	96	107	115	114	102	103	96	93	94
101	90	91	107	100	109	92	90	112	98
99	108	97	97	113	106	91	96	91	100
110	109	105	96	115	113	107	109	96	102
92	96	113	113	112	100	104	97	101	115
110	109	103	115	94	102	94	94	94	111
99	110	112	109	95	98	107	93	111	96
105	114	99	91	111	102	105	91	104	111
113	92	102	91	112	114	101	107	112	94
95	110	105	97	91	106	112	94	99	110
93	91	110	101	109	115	114	108	111	94
109	97	112	115	113	110	105	114	115	90
92	104	109	104	115	90				

Practice

Test whether the patients in the two groups have the same variance

```
> var.test(active.begin, placebo.begin)

         F test to compare two variances

data:  active.begin and placebo.end
F = 1.0068, num df = 192, denom df = 195, p-value = 0.9622
alternative hypothesis: true ratio of variances is not equal to 1
...

> var.test(active.begin, placebo.begin, ratio=0.5, "g")

         F test to compare two variances

data:  active.begin and placebo.end
F = 2.0136, num df = 192, denom df = 195, p-value = 7.252e-07
alternative hypothesis: true ratio of variances is greater than 0.5
...
```


PEKING UNIVERSITY
1898
非参数统计
讲义
孙山泽 编著
普通高等教育"十一五"国家级规划教材
高等学校统计学类系列教材
非参数
统计分析
王静龙 梁小筠 编著
1-book.cn
HIGHER EDUCATION PRESS

Sign Test

统计学方法及其应用 | 统计学基础 | 随机变量的函数

"A random variable is a quantity whose values are random and to which a probability distribution is assigned."

Counting statistics

Comparison function and sign function

$$\Psi^{(\theta)} = \psi(t-\theta) = \begin{cases} 1 & \text{if } t > \theta \\ 0 & \text{if } t \le \theta \end{cases}, \Psi^{(0)} = \psi(t) = \begin{cases} 1 & \text{if } t > 0 \\ 0 & \text{if } t \le 0 \end{cases}$$

Sign statistics

$$\Psi_i^{(0)} = \psi(X_i)$$

Sign sum statistic

$$B = \sum_{i=1}^{n} \Psi_i^{(0)} = \sum_{i=1}^{n} \psi(X_i)$$

Distribution of sign statistics

Distribution of sign statistics

For a random sample $X_1, \ldots, X_n$ that come from a population with cdf $F(x)$, if $F(0) = p$, then each sign statistic $\psi(X_i)$ has a Bernoulli$(1 - p)$ distribution.

Distribution of the sign sum statistic

For a random sample $X_1, \ldots, X_n$ from a population with cdf $F(x)$, if $F(0) = p$, then the sign sum statistic

$$B = \sum_{i=1}^{n} \psi(X_i)$$

has a Binomial$(n, 1 - p)$ distribution.

Tests of the median

Protocol

A random sample $X_1, \ldots X_n$ is observed from a population whose **median** is m. We like to test

(1) $H_0 : m = m_0$ versus $H_1 : m > m_0$ (assume $m \geq m_0$);

(2) $H_0 : m = m_0$ versus $H_1 : m < m_0$ (assume $m \leq m_0$);

(3) $H_0 : m = m_0$ versus $H_1 : m \neq m_0$;

Sign test for median, greater

Sign test for median, greater

Hypothesis:

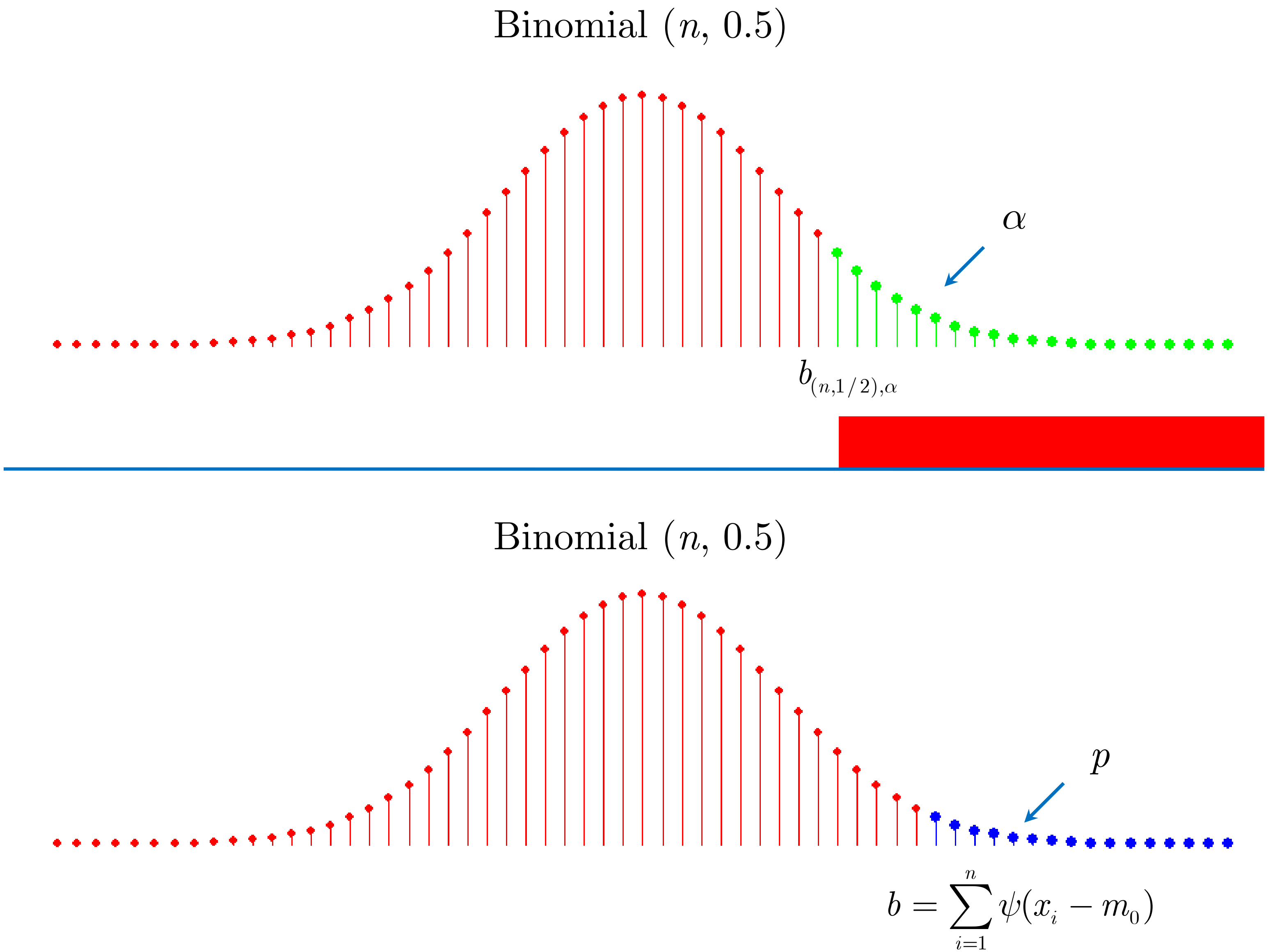
$$H_0 : m = m_0 \quad \text{versus} \quad H_1 : m > m_0$$

Test statistic:

$$B = \sum_{i=1}^{n} \psi(X_i - m_0), \ \#\left\{X_i > m_0\right\}$$

p-value:

$$p(\mathbf{x}) = P\left(B_{n,1/2} \geq \sum_{i=1}^{n} \psi(x_i - m_0)\right)$$


Binomial $(n,\, 0.5)$
α
$b_{(n,1/2),\alpha}$
Binomial $(n,\, 0.5)$
p
$b = \sum_{i=1}^{n} \psi(x_i - m_0)$

Sign test for median, less

Sign test for median, less

Hypothesis:

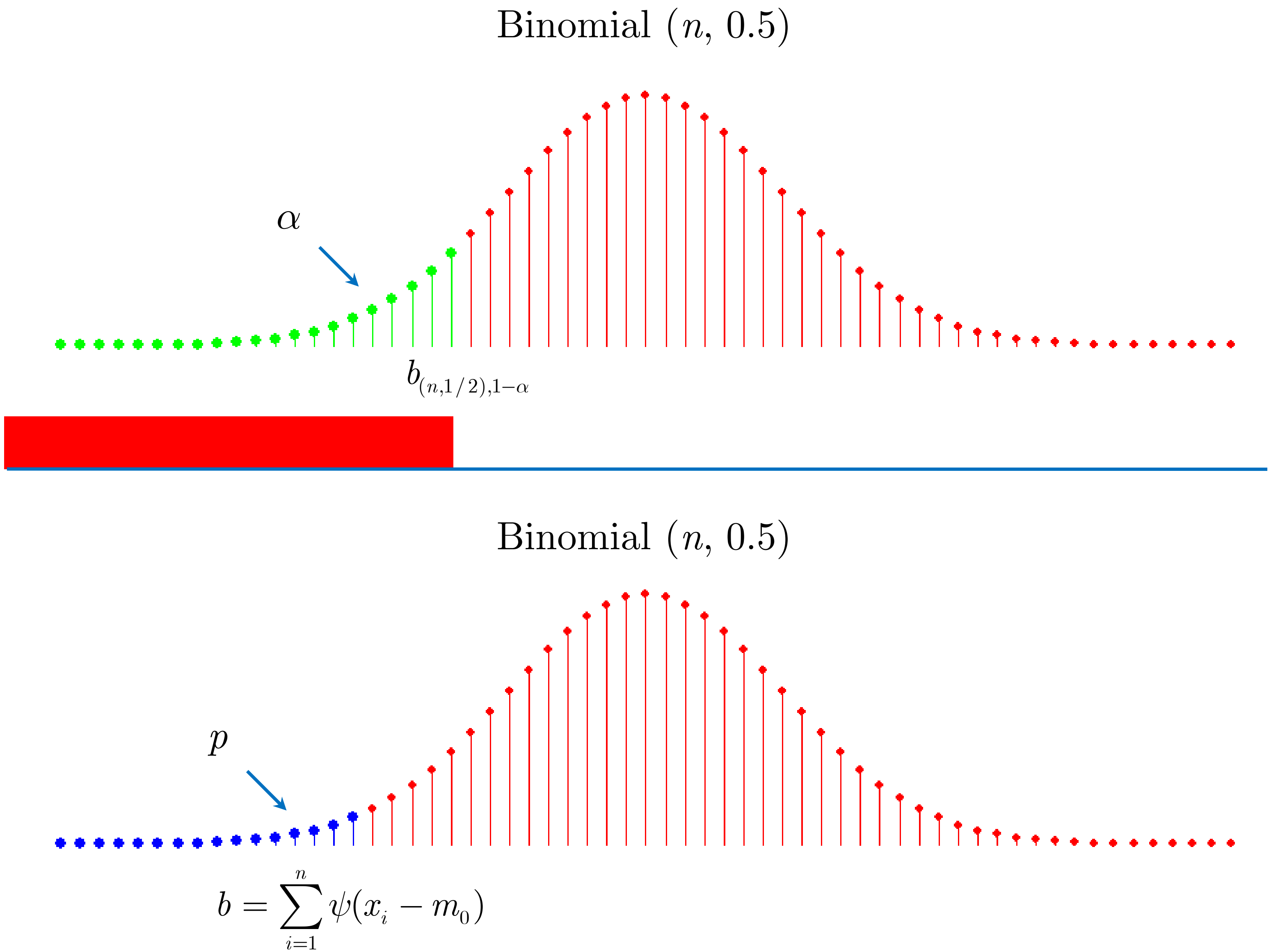
$$H_0 : m = m_0 \quad \text{versus} \quad H_1 : m < m_0$$

Test statistic:

$$B = \sum_{i=1}^{n} \psi(X_i - m_0),\ \#\left\{X_i > m_0\right\}$$

p-value:

$$p(\mathbf{x}) = P\left(B_{n,1/2} \le \sum_{i=1}^{n} \psi(x_i - m_0)\right)$$



Binomial $(n, 0.5)$
α
$b_{(n,1/2),1-\alpha}$
Binomial $(n, 0.5)$
p
$b = \sum_{i=1}^{n} \psi(x_i - m_0)$

Sign test for median, two-sided

Sign test for median, two-sided

Hypothesis:

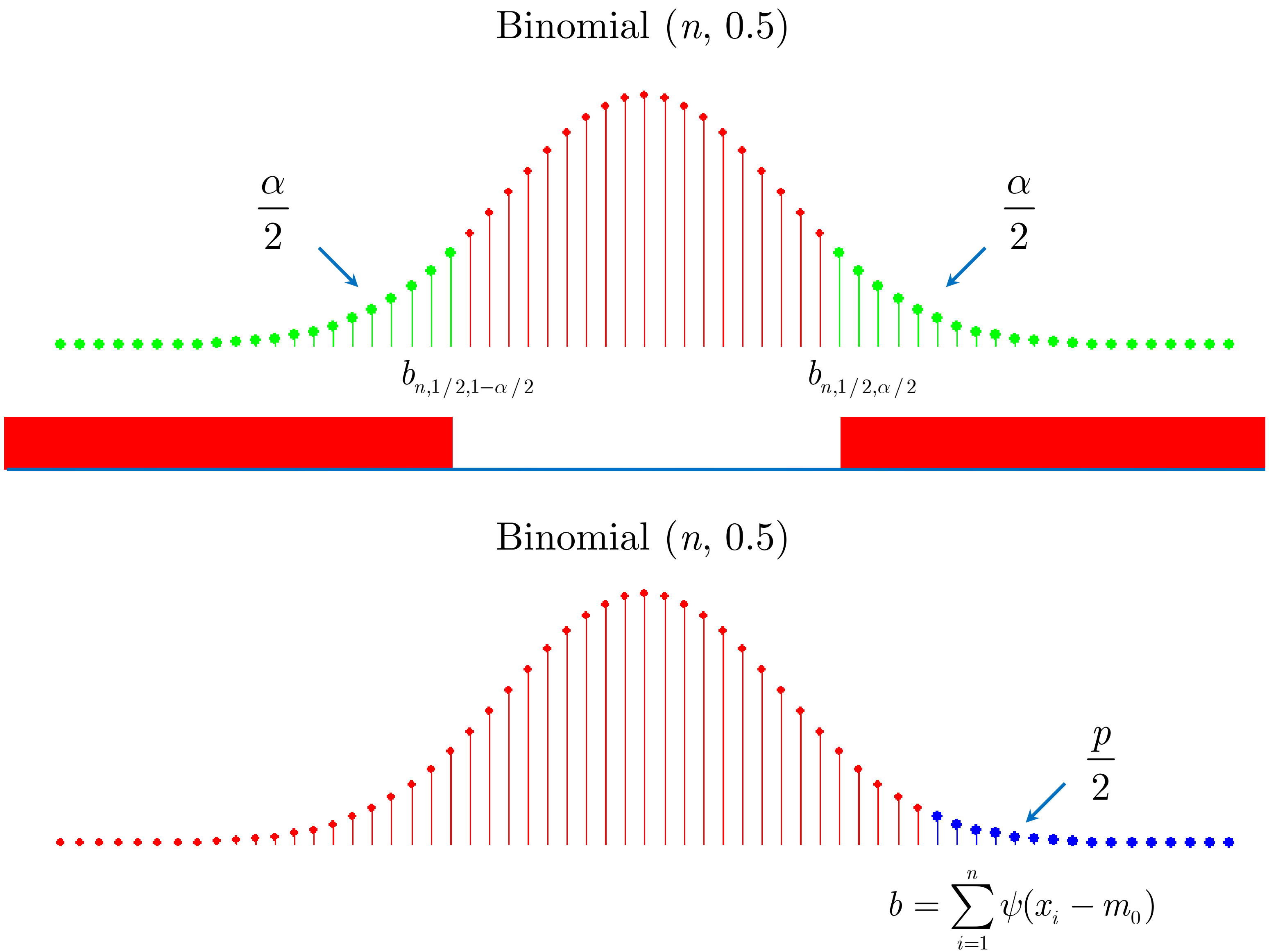
$$H_0 : m = m_0 \quad \text{versus} \quad H_1 : m \neq m_0$$

Test statistic:

$$B = \sum_{i=1}^{n} \psi(X_i - m_0),\ \#\left\{X_i > m_0\right\}$$

p-value:

$$p(\mathbf{x}) = \begin{cases} 2P\left(B_{n,1/2} \geq \sum_{i=1}^{n} \psi(x_i - m_0)\right), & \sum_{i=1}^{n} \psi(x_i - m_0) \geq \dfrac{n}{2} \\ 2P\left(B_{n,1/2} \leq \sum_{i=1}^{n} \psi(x_i - m_0)\right), & \sum_{i=1}^{n} \psi(x_i - m_0) < \dfrac{n}{2} \end{cases}$$



Binomial $(n,\ 0.5)$
$\frac{\alpha}{2}$
$\frac{\alpha}{2}$
$b_{n,1/2,1-\alpha/2}$
$b_{n,1/2,\alpha/2}$
Binomial $(n,\ 0.5)$
$\frac{p}{2}$
$b = \sum_{i=1}^{n} \psi(x_i - m_0)$

Sign of the change (DBP, Active group)

-23	-11	-7	-13	4	-32	-20	-18	7	-32
-14	-18	6	10	-4	-15	-21	-10	10	-20
-15	-10	-11	-10	0	-13	-14	-6	9	-19
-10	-19	-6	5	-17	-6	-15	6	-8	-19
-16	2	-6	-14	-22	-11	-23	-6	-12	-12
0	0	-3	-14	-34	-6	-30	-17	-17	-1
-30	-31	-17	-8	8	-23	-12	9	-33	4
-18	-16	-28	-10	-8	-20	-8	-20	-8	19
-12	-11	-19	-12	-10	-20	-11	-2	-17	-24
-18	-13	25	4	-13	-1	-7	-2	-22	-25
-19	-8	-17	-10	-27	-1	-6	-19	4	-16
-29	4	-8	-16	-16	1	-7	-31	-9	0
-4	-16	-5	-6	-14	-3	0	31	-10	-23
-14	-24	-11	-2	20	-5	-21	-1	-2	-3
-21	-5	-10	-12	0	-5	10	-26	-9	-10
16	-15	-26	1	-18	-19	-16	10	0	4
-9	-4	1	-7	-5	-11	-8	-5	-19	-5
-2	0	-18	-33	-28	-23	-18	-13	-8	-3
-2	1	6	-20	-11	-6	-2	-32	-27	-22
-17	-12	-7							

Sign of the change (DBP, Placebo group)

-15	-2	6	-9	-3	-1	9	15	17	3
10	-10	-18	2	-18	5	-3	-1	10	5
20	-14	-6	3	10	-3	1	16	0	0
-5	6	-14	1	8	-9	10	6	-10	4
-3	-12	5	17	2	20	3	-8	-16	-12
-13	-6	5	17	8	-5	-8	-9	12	-7
-21	-10	-2	-4	-18	6	-8	-8	10	20
5	-11	-19	-13	19	-6	-10	-2	8	2
-13	11	8	0	-17	0	-16	-23	20	-21
-6	4	-5	-12	-17	-20	-6	2	-5	5
2	0	15	1	4	18	15	19	-17	-5
-10	-14	8	4	20	-1	10	12	-5	5
-5	20	7	7	-4	0	-2	1	0	1
0	0	0	-10	7	4	0	0	-26	2
6	20	23	-2	11	-4	-7	20	4	3
20	23	-2	11	-4	-7	20	4	3	20
15	2	-10	2	-16	-1	-6	3	-19	7
6	7	2	22	4	-6	20	-35	-3	23
-7	-3	5	10	15	-8	32	27	22	17
12	7	2	-3	-8	-13				

Practice

Test whether the patients in the same group have lower blood pressure

```
> binom.test(sum(active.diff>0),193, 0.5, alternative="l")

         Exact binomial test

data:  sum(active.diff > 0) and 193
number of successes = 28, number of trials = 193, p-value < 2.2e-16
alternative hypothesis: true probability of success is less than 0.5
...

> binom.test(sum(placebo.diff>0),196, 0.5, alternative="l")

         Exact binomial test

data:  sum(placebo.diff > 0) and 196
number of successes = 100, number of trials = 196, p-value = 0.6395
alternative hypothesis: true probability of success is less than 0.5
...
```

Compare it with one-sample t test!

Signed Rank Test

"A random variable is a quantity whose values are random and to which a probability distribution is assigned."

Rank statistics

Rank statistics

For a random sample $X_1,\ldots,X_n$, the **rank statistic** of the i-th random variable, R_i, is the position of X_i in the order statistics $X_{(1)},\ldots,X_{(n)}$, in other words

$$R_i = r, \text{ if } X_i = X_{(r)}$$

Distribution of rank statistics

Marginal distributions of rank statistics

For a random sample $X_1, \ldots, X_n$, each rank statistic R_i has a discrete uniform distribution

$$P(R_i = r) = \frac{1}{n}, \quad r = 1, \ldots, n$$

Joint distribution of all rank statistic

For a random sample $X_1, \ldots, X_n$, let $\mathbf{R} = (R_1, \ldots, R_n)$ be the n rank statistics. Then $\mathbf{R}$ has a uniform distribution

$$P(\mathbf{R} = \mathbf{r}) = \frac{1}{n!},$$

where $\mathbf{r} = (r_1, \ldots, r_n)$ is a permutation of $(1, \ldots, n)$.

Signed rank statistics

Signed rank statistics

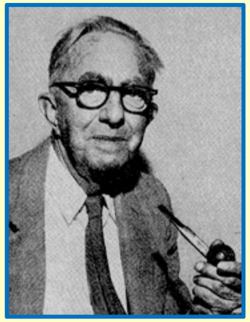
For a random sample $X_1,\ldots,X_n$, the **absolute rank statistic** of X_i, R_i^+, is the rank statistic of $|X_i|$ in $|X_1|,\ldots,|X_n|$, and the **signed rank statistic** of X_i is

$$\psi(X_i)R_i^+.$$

The **Wilcoxon signed rank statistic** is

$$W^+ = \sum_{i=1}^{n} \psi(X_i)R_i^+.$$

Frank Wilcoxon



Properties

W^+ has $n(n+1)/2+1$ distinct values,

$$0, 1, \ldots, n(n+1)\,/\,2$$

The minimum is

$$0 = 0 \times 1 + 0 \times 2 + \cdots + 0 \times n$$

The maximum is

$$n(n+1)\,/\,2 = 1 \times 1 + 1 \times 2 + \cdots + 1 \times n$$

Therefore,

W^+ has a discrete distribution

Distribution

For a random sample $X_1, \ldots, X_n$, the Wilcoxon signed rank statistic $W^+ = \sum_{i=1}^{n} \psi(X_i) R_i^+$ has a

$$P(W^+ = w^+) = \frac{C_n(w^+)}{2^n}$$

distribution, where $C_n(w^+)$ is the number of subsets of n numbers $(1, 2, \ldots, n)$ that sum up to w^+.

Recurrence

$$C_n(w^+) \;=\; C_{n-1}(w^+ - n) \;+\; C_{n-1}(w^+)$$

If n is selected, we need to further select n–1 numbers from 1, …, n–1, and the sum of these n–1 numbers should be $w^+–n$.

If n is not selected, we need to select n–1 numbers from 1, …, n–1, and the sum of these n numbers should still be w^+.

$$C_1(0) = 1;$$
$$C_1(1) = 1;$$
$$C_1(w^+) = 0 \text{ if } w \neq 0,1.$$

Symmetry

Suppose the sum of some selected numbers is w^+. Then the sum of the rest numbers should be $n(n+1)/2 - w^+$.

That is to say, the number of subsets of n numbers $(1,2,\ldots,n)$ that sum up to w^+ is equal to the number of subsets of n numbers that sum up to $n(n+1)/2 - w^+$.

$$C_n(w^+) = C_n(n(n+1)/2 - w^+)$$

Therefore,

$$P(W^+ = w^+) = P(W^+ = n(n+1)/2 - w^+)$$

Hence,

$$P(W^+ = n(n+1)/4 - w^+) = P(W^+ = n(n+1)/4 + w^+)$$

The distribution of W^+ is symmetric about $n(n+1)/4$.

Distribution functions

Density, distribution function, quantile function and random generation for the distribution of the Wilcoxon Signed Rank statistic obtained from a sample with size n.

Usage:

```
dsignrank(x, n, log = FALSE)
psignrank(q, n, lower.tail = TRUE, log.p = FALSE)
qsignrank(p, n, lower.tail = TRUE, log.p = FALSE)
rsignrank(nn, n)
```

Tests of the symmetric point

A random sample $X_1, \dots X_n$ is observed from a population that is **symmetric** about m. We like to test

(1) $H_0 : m = m_0$ versus $H_1 : m > m_0$;

(2) $H_0 : m = m_0$ versus $H_1 : m < m_0$;

(3) $H_0 : m = m_0$ versus $H_1 : m \neq m_0$.

Signed rank test, greater

Hypothesis:

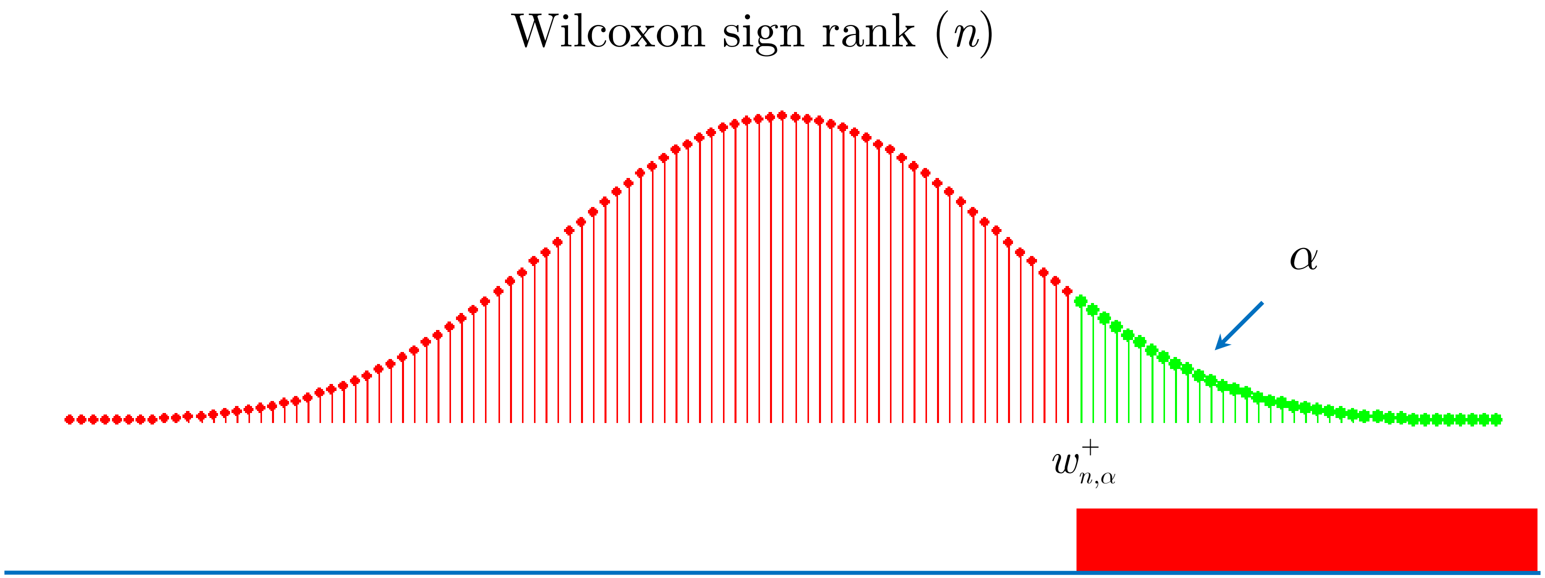
$$H_0 : m = m_0 \quad \text{versus} \quad H_1 : m > m_0$$

Test statistic:

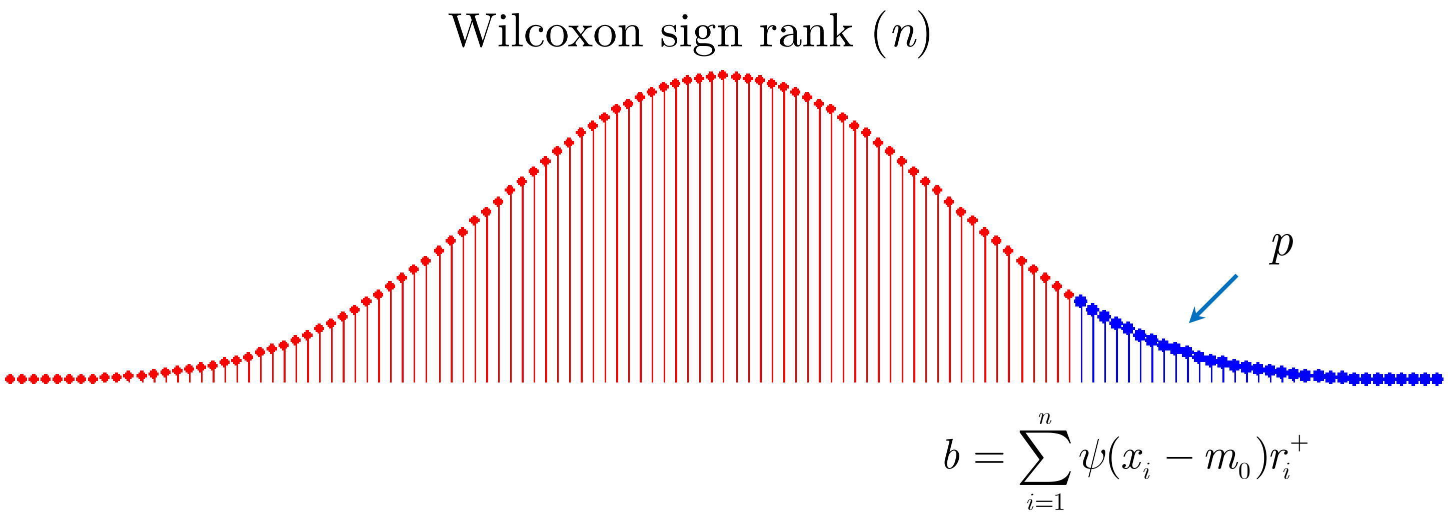
$$W^+ = \sum_{i=1}^{n} \psi(X_i - m_0) R_i^+$$

p-value:

$$p(\mathbf{x}) = P\left(W_n^+ \geq \sum_{i=1}^{n} \psi(x_i - m_0) r_i^+ \right)$$



Wilcoxon sign rank (n)
α
$w^{+}_{n,\alpha}$



Wilcoxon sign rank (n)
p
$b = \sum_{i=1}^{n} \psi(x_i - m_0) r_i^{+}$

Signed rank test, less

Hypothesis:

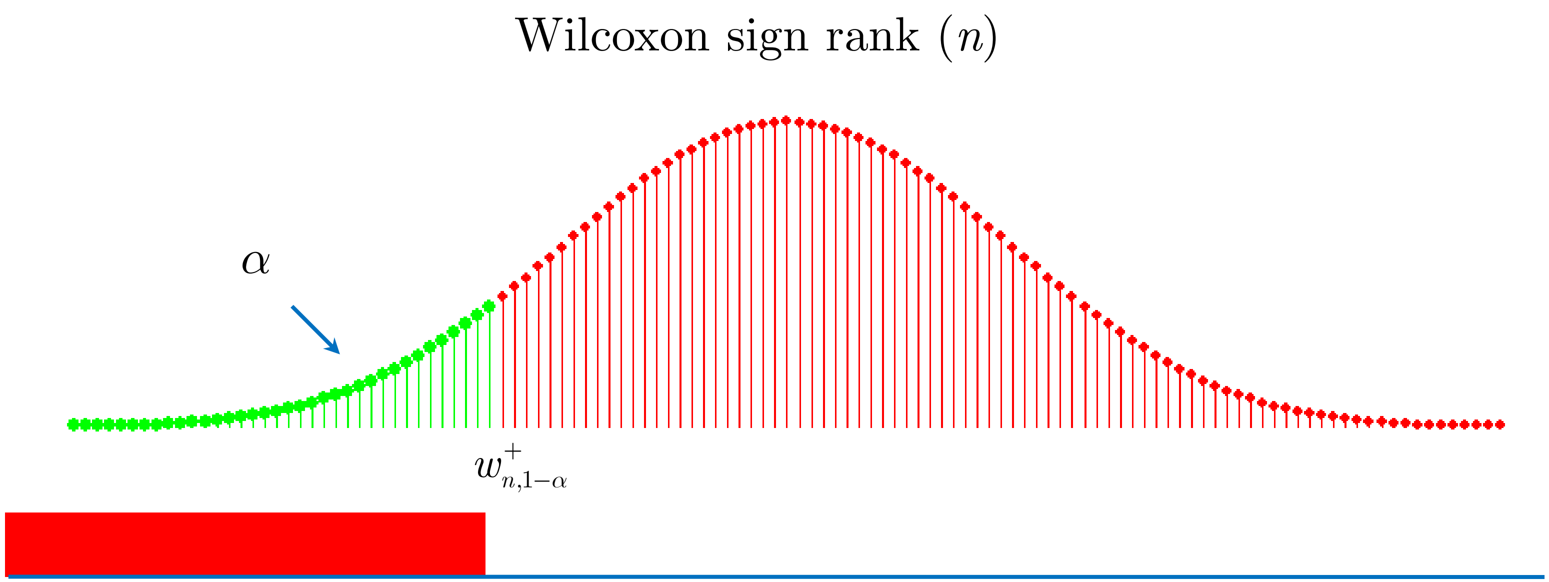
$$H_0 : m = m_0 \quad \text{versus} \quad H_1 : m < m_0$$

Test statistic:

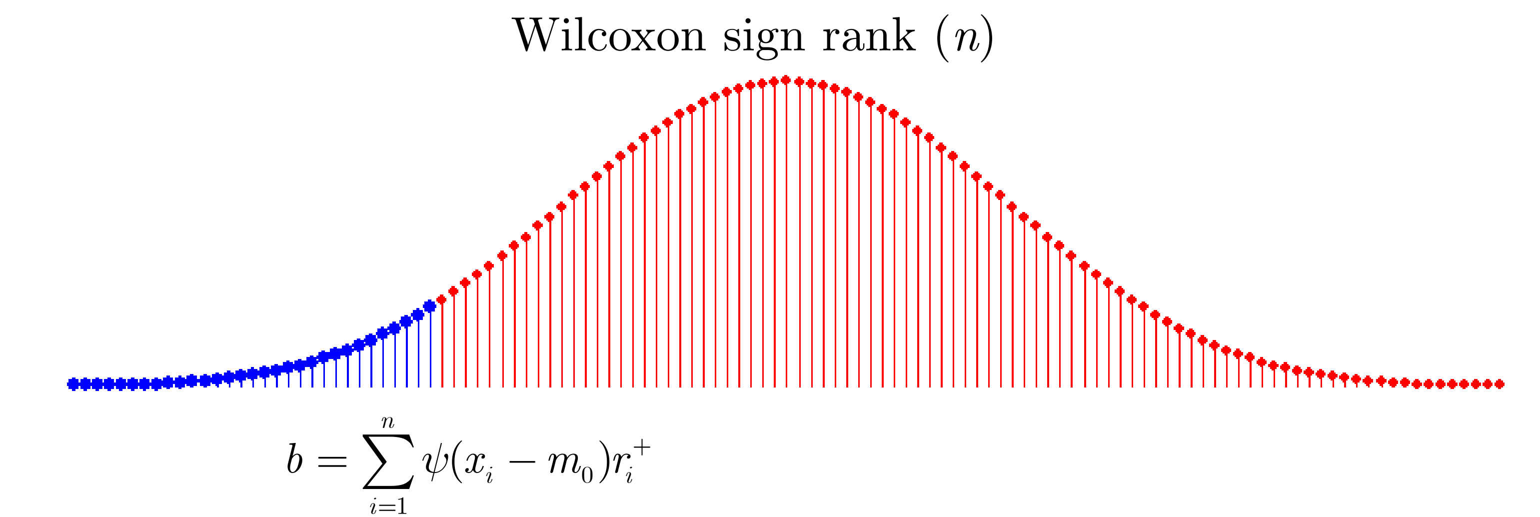
$$W^+ = \sum_{i=1}^{n} \psi(X_i - m_0) R_i^+$$

p-value:

$$p(\mathbf{x}) = P\left(W_n^+ \leq \sum_{i=1}^{n} \psi(x_i - m_0) r_i^+ \right)$$



Wilcoxon sign rank (n)
α
$w^{+}_{n,1-\alpha}$



Wilcoxon sign rank (n)
$b = \sum_{i=1}^{n} \psi(x_i - m_0) r_i^{+}$

Signed rank test, two-sided

Hypothesis:

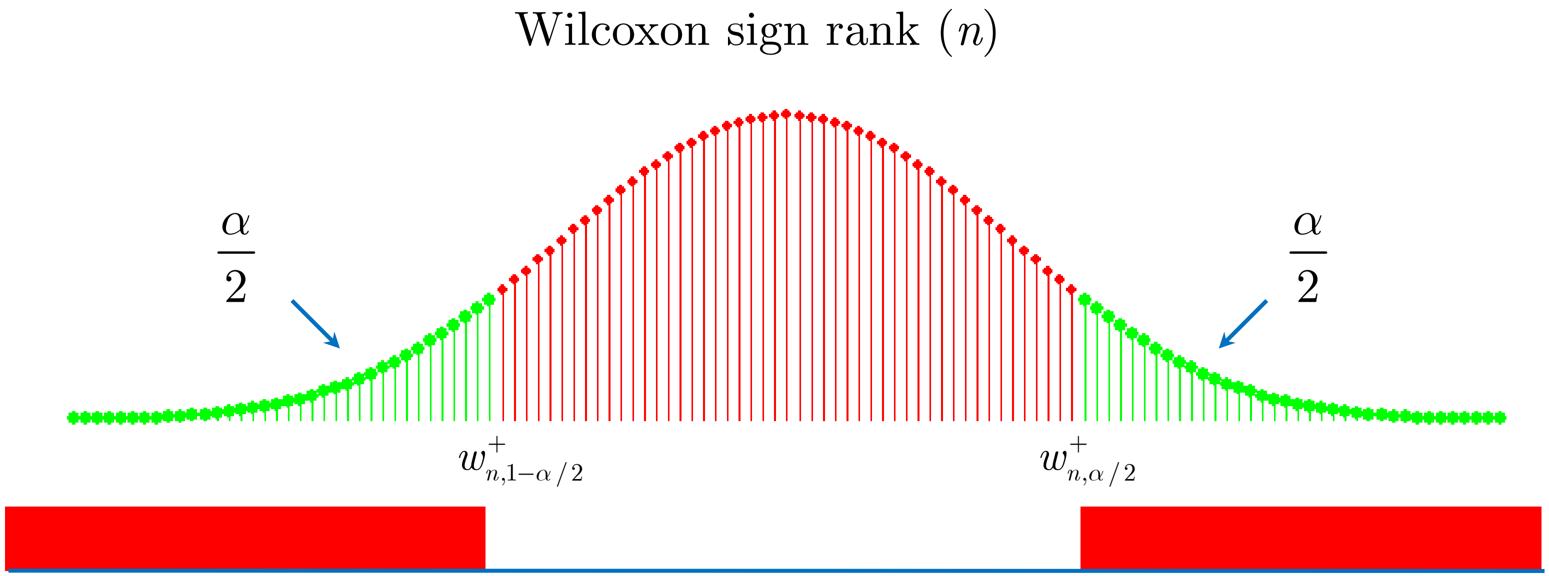
$$H_0 : m = m_0 \quad \text{versus} \quad H_1 : m \neq m_0$$

Test statistic:

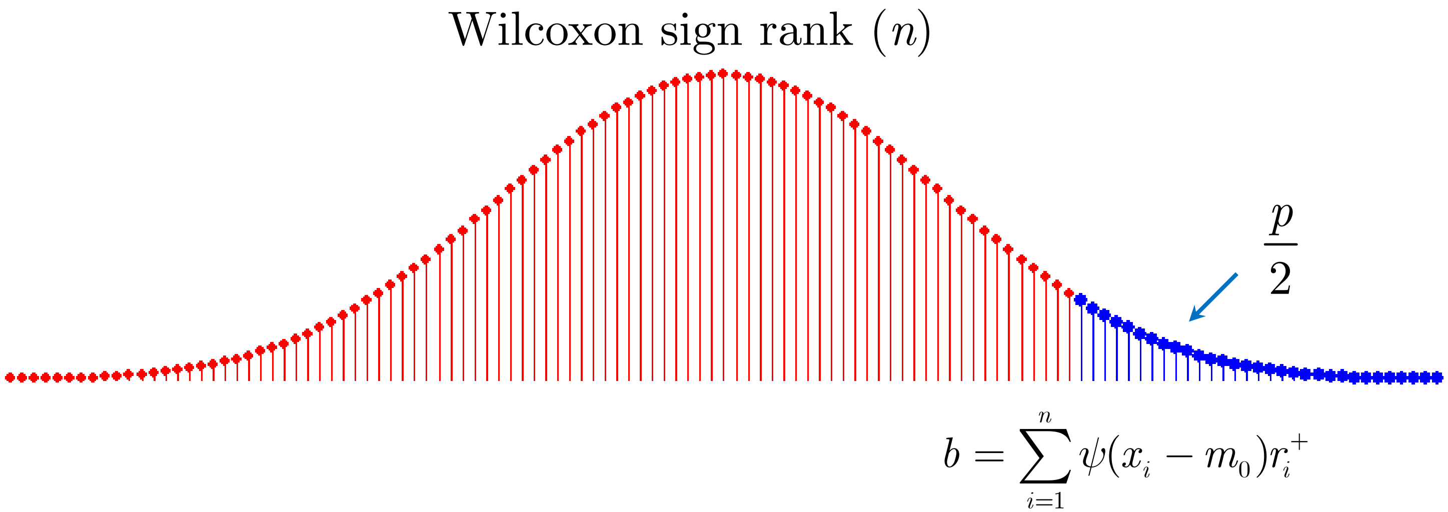
$$W^+ = \sum_{i=1}^{n} \psi(X_i - m_0) R_i^+$$

p-value:

$$p(\mathbf{x}) = \begin{cases} 2P\left(W_n^+ \geq \sum_{i=1}^{n} \psi(x_i - m_0) r_i^+\right), & \sum_{i=1}^{n} \psi(x_i - m_0) r_i^+ \geq \dfrac{n(n+1)}{4} \\ 2P\left(W_n^+ \leq \sum_{i=1}^{n} \psi(x_i - m_0) r_i^+\right), & \sum_{i=1}^{n} \psi(x_i - m_0) r_i^+ < \dfrac{n(n+1)}{4} \end{cases}$$



Wilcoxon sign rank (n)
$\frac{\alpha}{2}$
$\frac{\alpha}{2}$
$w^+_{n,1-\alpha/2}$
$w^+_{n,\alpha/2}$



Wilcoxon sign rank (n)
$\frac{p}{2}$
$b = \sum_{i=1}^{n} \psi(x_i - m_0) r_i^+$

Rank Sum Test

统计学方法及其应用 > 统计学基础 > 随机变量的函数

“A random variable is a quantity whose values are random and to which a probability distribution is assigned.”

Rank sum statistic

For **two** random samples $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$, the **Wilcoxon rank sum statistic** is

$$W_Y = \sum_{i=1}^{n} R_Y^i,$$

where R_Y^i is the rank statistic of Y_i in the mixed sample $X_1, \ldots, X_m, Y_1, \ldots, Y_n$.

Properties

W_Y has $mn + 1$ distinct values,

$$n(n+1)\,/\,2, n(n+1)\,/\,2+1, \ldots, n(n+1)\,/\,2+mn$$

The minimum is

$$n(n+1)\,/\,2 = 1+2+\cdots+n$$

The maximum is

$$n(n+1)\,/\,2+mn = (m+1)+(m+2)+\cdots+(m+n)$$

Therefore,

W_Y has a discrete distribution

Distribution

If two random samples $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$ come from the same population, then the Wilcoxon rank sum statistic $W_Y = \sum_{i=1}^{n} R_Y^i$, has a

$$P(W_Y = w) = \frac{m!n!t_{m,n}(w)}{(m+n)!} = \frac{t_{m,n}(w)}{\begin{pmatrix} m+n \\ n \end{pmatrix}}$$

distribution, where, $w = n(n+1)/2, \ldots, n(n+1)/2 + mn$.

$t_{m,n}(w)$ is the number of ways of choosing n numbers $k_1, \ldots, k_n$ from $1, \ldots, m+n$ such that $\sum_{i=1}^{n} k_i = w$.

Recurrence

$$t_{m,n}(w) \;=\; t_{m,n-1}(w-m-n) \;+\; t_{m-1,n}(w)$$

If $m+n$ is selected, we need to further select n–1 numbers from 1, …, $m+n$–1, and the sum of these n–1 numbers should be w–m–n.

If $m+n$ is not selected, we need to select n numbers from 1, …, $m+n$–1, and the sum of these n numbers should still be w.

$$t_{i,0}(0) = 1 \;(i = 1, \ldots, m);$$
$$t_{i,0}(w) = 0 \;(i = 1, \ldots, m, \text{if } w \neq 0);$$
$$t_{0,j}\left(\tfrac{j(j+1)}{2}\right) = 1 \;(j = 1, \ldots, n);$$
$$t_{0,j}(w) = 0 \;(j = 1, \ldots, m, \text{ if } w \neq \tfrac{j(j+1)}{2})$$

Symmetry

Suppose the n selected numbers are $a_1,\dots,a_n$ with the sum $\Sigma_{i=1}^{n} a_i = w$.
Let $b_i = (m+n+1) - a_i$ $(i = 1,\dots,n)$, then $\Sigma_{i=1}^{n} b_i = n(m+n+1) - w$.
That is to say, the number of selecting n numbers with the sum w is equal to the number of selecting n numbers with the sum $n(m+n+1) - w$.

$$t_{m,n}(w) = t_{m,n}(n(m+n+1) - w)$$

Therefore,

$$P(W_Y = w) = P(W_Y = n(m+n+1) - w)$$

Hence,

$$P(W_Y = n(m+n+1)/2 - w) = P(W_Y = n(m+n+1)/2 + w)$$

The distribution of W_Y is symmetric about $n(m+n+1)/2$.

The Mann-Whitney statistic

For **two** random samples $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$, the **Mann − Whitney statistic** is

$$U = \sum_{i=1}^{m} \sum_{j=1}^{n} \psi(Y_j - X_i),$$

and

$$W = U + \frac{n(n+1)}{2}$$

Distribution functions

Density, distribution function, quantile function and random generation for the distribution of the Wilcoxon rank sum statistic obtained from samples with size ‘m’ and ‘n’, respectively.

Usage:

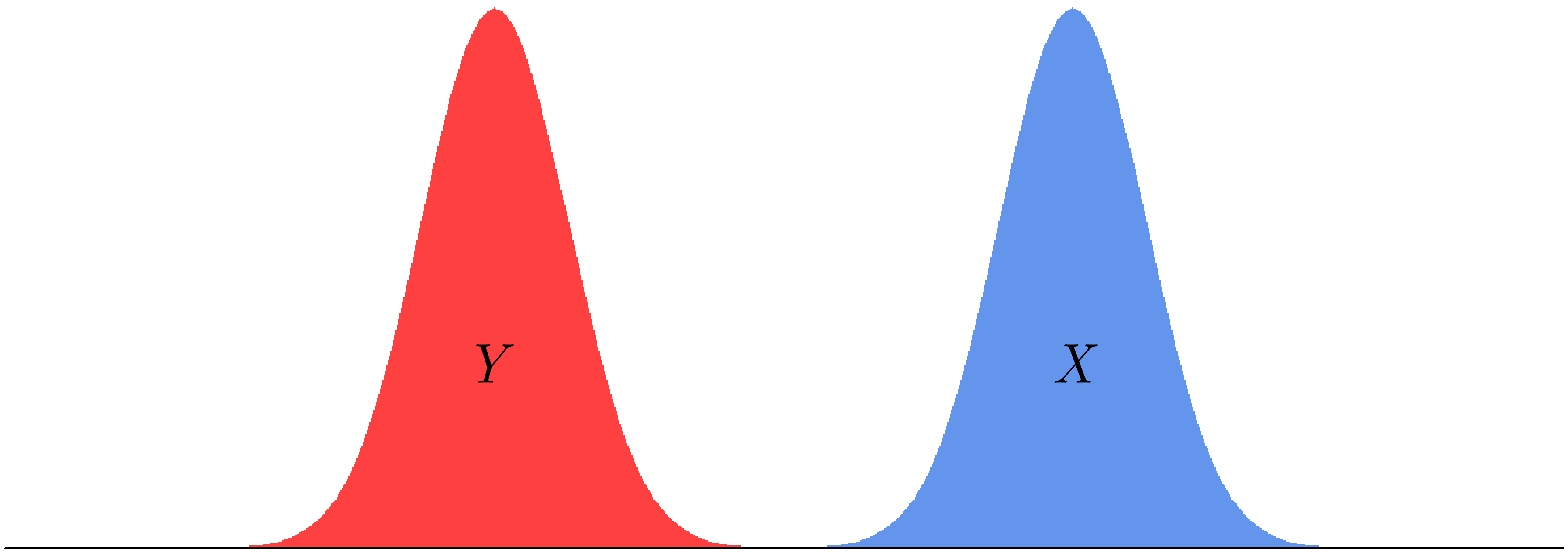
```
dwilcox(x, m, n, log = FALSE)
pwilcox(q, m, n, lower.tail = TRUE, log.p = FALSE)
qwilcox(p, m, n, lower.tail = TRUE, log.p = FALSE)
rwilcox(nn, m, n)
```

Comparison of two distributions

- Null hypothesis $H_0 : X \overset{d}{\rightleftharpoons} Y$

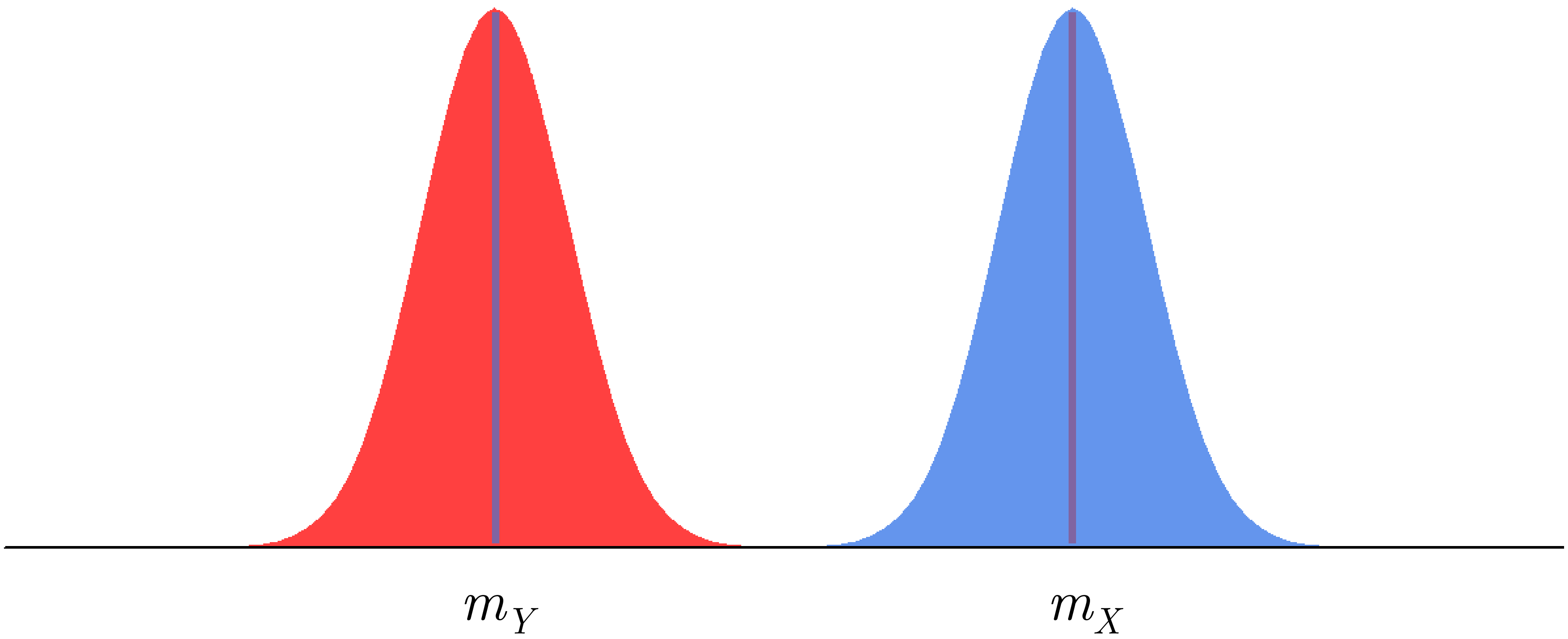
Comparison of two distributions

$$P(X > Y) > 1/2 \text{ implies } "X > Y"$$



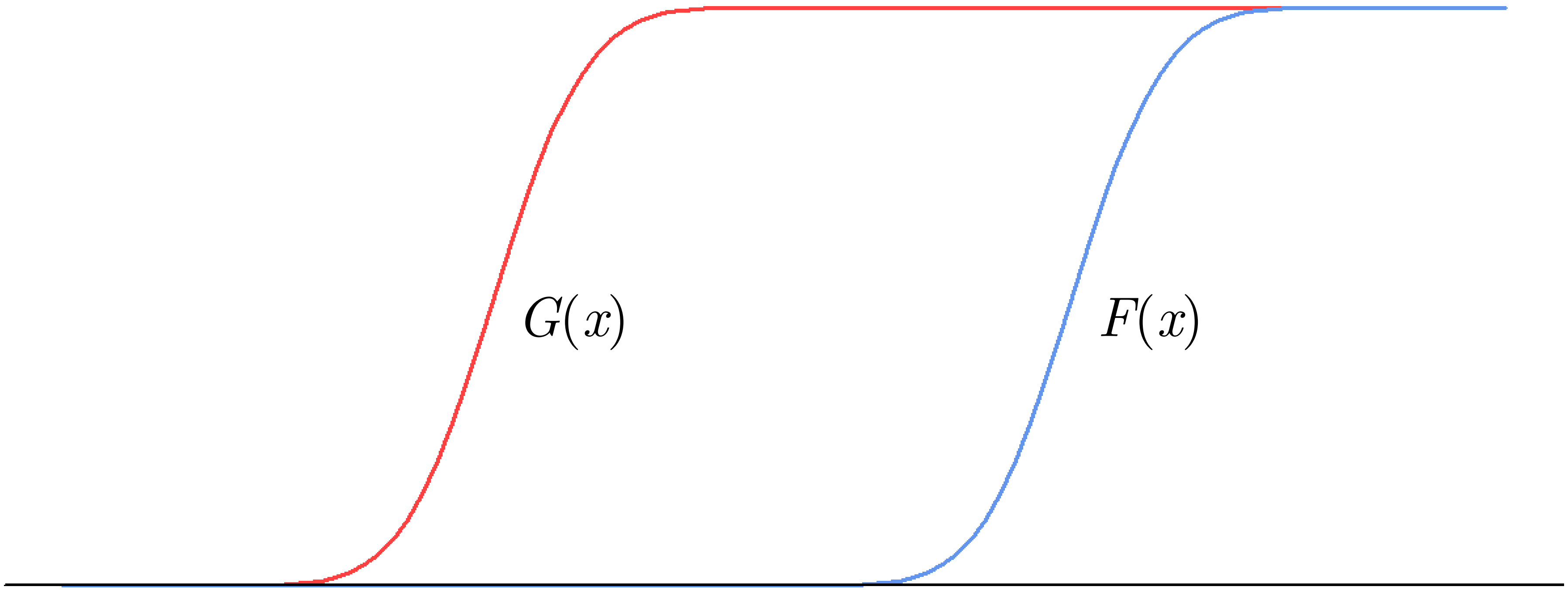
Comparison of two medians

$$m_X > m_Y \text{ implies } "X > Y"$$



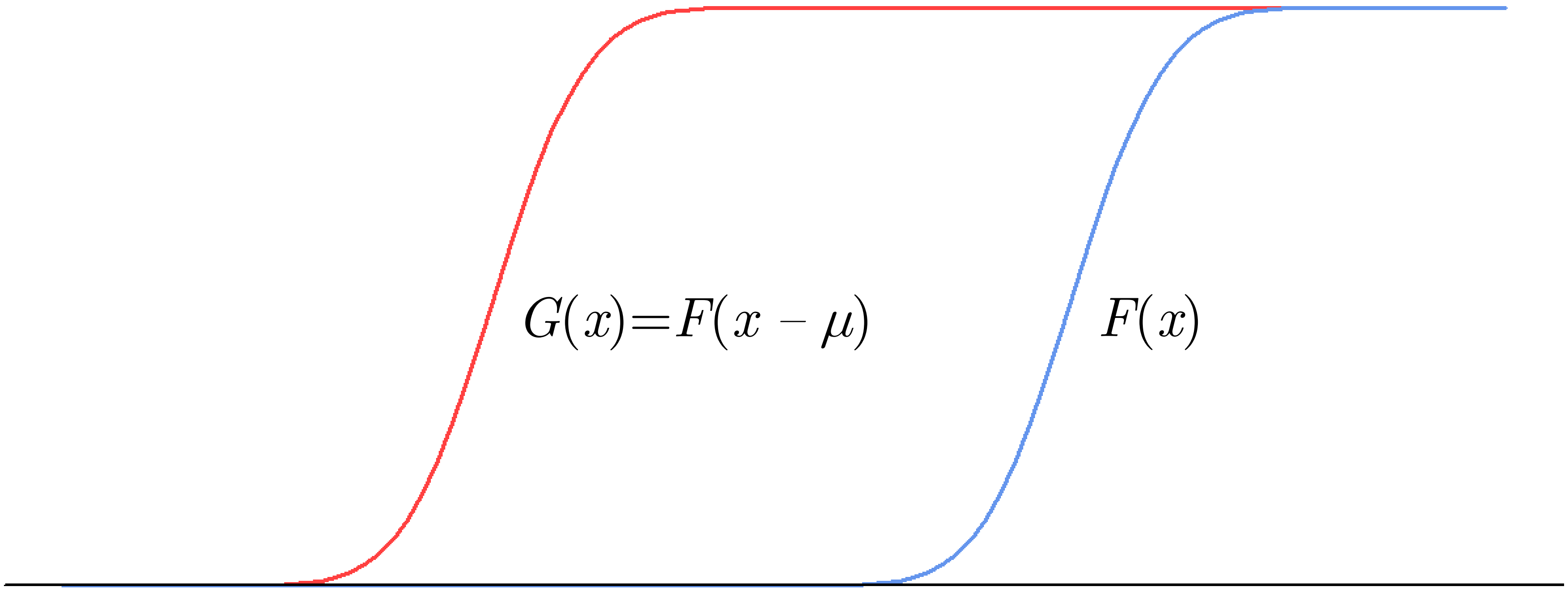
Comparison of two CDFs

$$F(x) < G(x) \text{ implies } "X > Y"$$



Comparison of location parameters

$$G(x) = F(x - \mu),\ \mu < 0 \text{ implies } F(x) < G(x)$$



Tests of identical distributions

Two random samples $X_1, \ldots X_m$ and $Y_1, \ldots Y_n$ are observed from two populations X and Y. We like to test

$$(1)\quad H_0 : X \overset{d}{\rightleftharpoons} Y \qquad \text{versus} \quad H_1 : P(X > Y) > 1/2;$$

$$(2)\quad H_0 : X \overset{d}{\rightleftharpoons} Y \qquad \text{versus} \quad H_1 : P(X > Y) < 1/2;$$

$$(3)\quad H_0 : X \overset{d}{\rightleftharpoons} Y \qquad \text{versus} \quad H_1 : P(X > Y) \neq 1/2.$$

Here,

$X \overset{d}{\rightleftharpoons} Y$ means X and Y are identically distributed.

$P(X > Y) > 1/2$ implies "$X > Y$"

Tests of the median

Two random samples $X_1, \dots X_m$ and $Y_1, \dots Y_n$ are observed from two populations X and Y, with **medians** being m_X and m_Y, respectively. We like to test

(1) $H_0 : X \overset{d}{\Longleftrightarrow} Y$ versus $H_1 : m_X > m_Y$;

(2) $H_0 : X \overset{d}{\Longleftrightarrow} Y$ versus $H_1 : m_X < m_Y$;

(3) $H_0 : X \overset{d}{\Longleftrightarrow} Y$ versus $H_1 : m_X \neq m_Y$.

Here,

$X \overset{d}{\Longleftrightarrow} Y$ means X and Y are identically distributed.

$m_X > m_Y$ implies "$X > Y$".

Tests of CDFs

Two random samples $X_1, \ldots X_m$ and $Y_1, \ldots Y_n$ are observed from two populations X and Y, with cdf $F(x)$ and $G(x)$, respectively. We like to test

(1) $H_0 : X \overset{d}{\Longleftrightarrow} Y$ versus $H_1 : F(x) < G(x)$ for $\forall x$;

(2) $H_0 : X \overset{d}{\Longleftrightarrow} Y$ versus $H_1 : F(x) > G(x)$ for $\forall x$;

(3) $H_0 : X \overset{d}{\Longleftrightarrow} Y$ versus $H_1 : F(x) < G(x)$ or $F(x) > G(x)$ for $\forall x$.

Here,

$X \overset{d}{\Longleftrightarrow} Y$ means X and Y are identically distributed.

$F(x) < G(x)$ implies "$X > Y$".

Tests of location parameter

Two random samples $X_1, \ldots X_m$ and $Y_1, \ldots Y_n$ are observed from two populations X and Y, with cdf $F(x)$ and $G(x) = F(x - \mu)$, respectively. We like to test

(1) $H_0 : X \overset{d}{\Longleftrightarrow} Y$ versus $H_1 : \mu < 0$;

(2) $H_0 : X \overset{d}{\Longleftrightarrow} Y$ versus $H_1 : \mu > 0$;

(3) $H_0 : X \overset{d}{\Longleftrightarrow} Y$ versus $H_1 : \mu \neq 0$.

Here,

$X \overset{d}{\Longleftrightarrow} Y$ means X and Y are identically distributed.

$\mu < 0$ implies $F(x) < G(x) \Rightarrow$ implies "$X > Y$".

Wilcoxon rank sum test, one sided

Hypothesis:

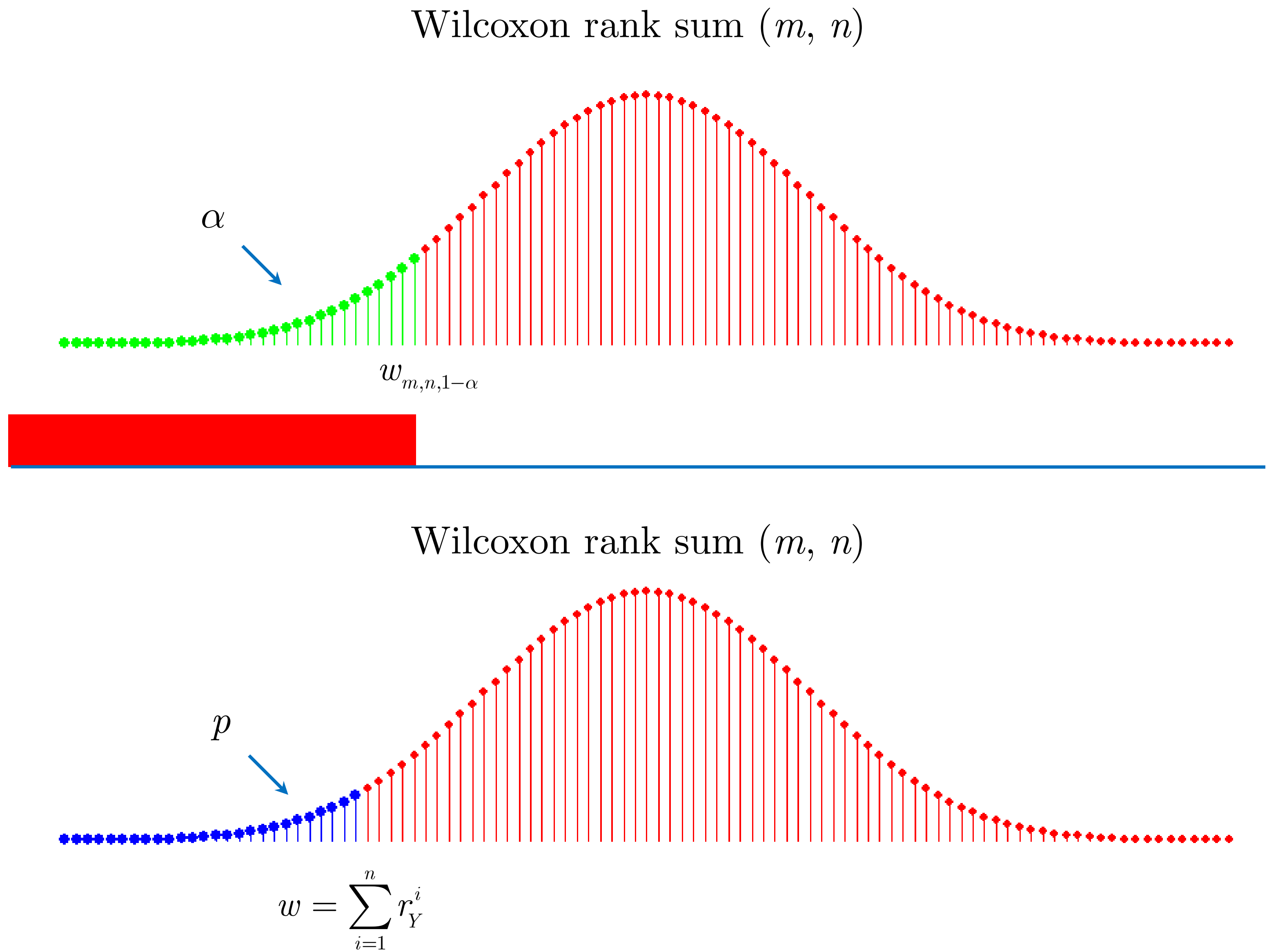
$$H_0 : X \xLeftrightarrow{d} Y \text{ versus } H_1 : P(X > Y) > 1/2;$$
$$H_0 : X \xLeftrightarrow{d} Y \text{ versus } H_1 : F(x) < G(x) \text{ for } \forall x;$$
$$H_0 : X \xLeftrightarrow{d} Y \text{ versus } H_1 : m_X > m_Y;$$
$$H_0 : X \xLeftrightarrow{d} Y \text{ versus } H_1 : \mu < 0;$$

Test statistic:

$$W_Y = \sum_{i=1}^{n} R_Y^i$$

p-value:

$$p(\mathbf{x}) = P\left(W_Y \leq \sum_{i=1}^{n} r_y^i\right)$$


Wilcoxon rank sum $(m,\ n)$
α
$w_{m,n,1-\alpha}$
Wilcoxon rank sum $(m,\ n)$
p
$w = \sum_{i=1}^{n} r_Y^i$

Wilcoxon rank sum test, one sided

Hypothesis:

$$H_0 : X \overset{d}{\rightleftharpoons} Y \text{ versus } H_1 : P(X > Y) < 1/2;$$

$$H_0 : X \overset{d}{\rightleftharpoons} Y \text{ versus } H_1 : F(x) > G(x) \text{ for } \forall x;$$

$$H_0 : X \overset{d}{\rightleftharpoons} Y \text{ versus } H_1 : m_X < m_Y;$$

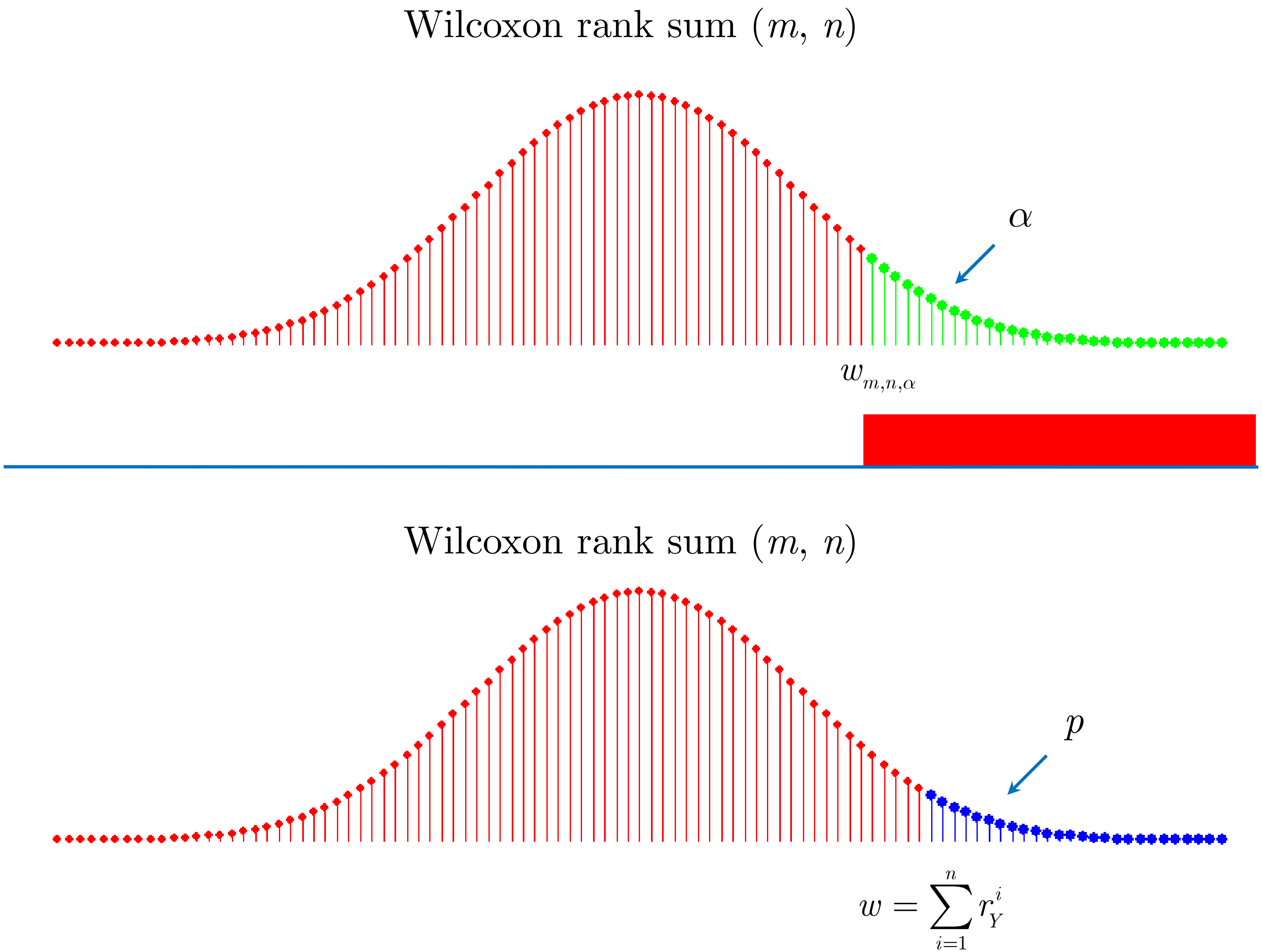
$$H_0 : X \overset{d}{\rightleftharpoons} Y \text{ versus } H_1 : \mu > 0;$$

Test statistic:

$$W_Y = \sum_{i=1}^{n} R_Y^i$$

p-value:

$$p(\mathbf{x}) = P\left(W_Y \le \sum_{i=1}^{n} r_y^i\right)$$



Wilcoxon rank sum $(m,\ n)$
α
$w_{m,n,\alpha}$
Wilcoxon rank sum $(m,\ n)$
p
$w = \sum_{i=1}^{n} r_Y^i$

Wilcoxon rank sum test, two-sided

Hypothesis:

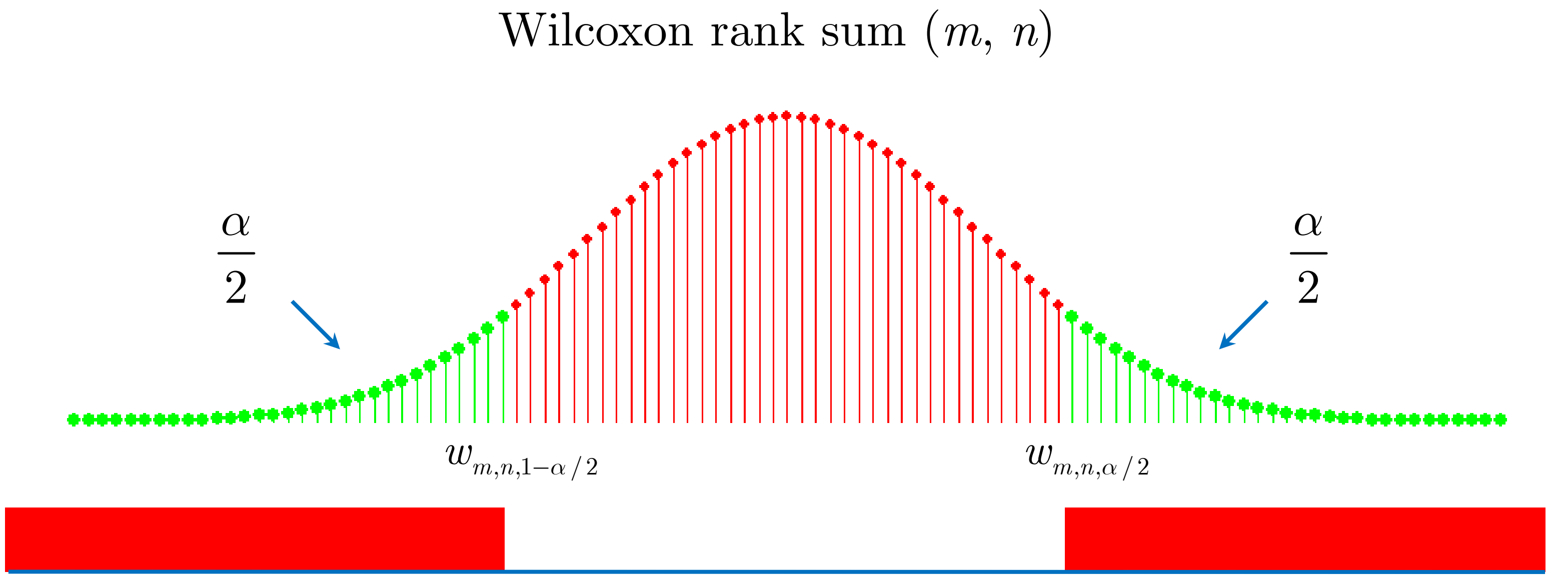
$$H_0 : X \overset{d}{\Longleftrightarrow} Y \text{ versus } H_1 : P(X > Y) \neq 1/2;$$
$$H_0 : X \overset{d}{\Longleftrightarrow} Y \text{ versus } H_1 : F(x) \neq G(x) \text{ for } \forall x;$$
$$H_0 : X \overset{d}{\Longleftrightarrow} Y \text{ versus } H_1 : m_X \neq m_Y;$$
$$H_0 : X \overset{d}{\Longleftrightarrow} Y \text{ versus } H_1 : \mu \neq 0;$$

Test statistic:

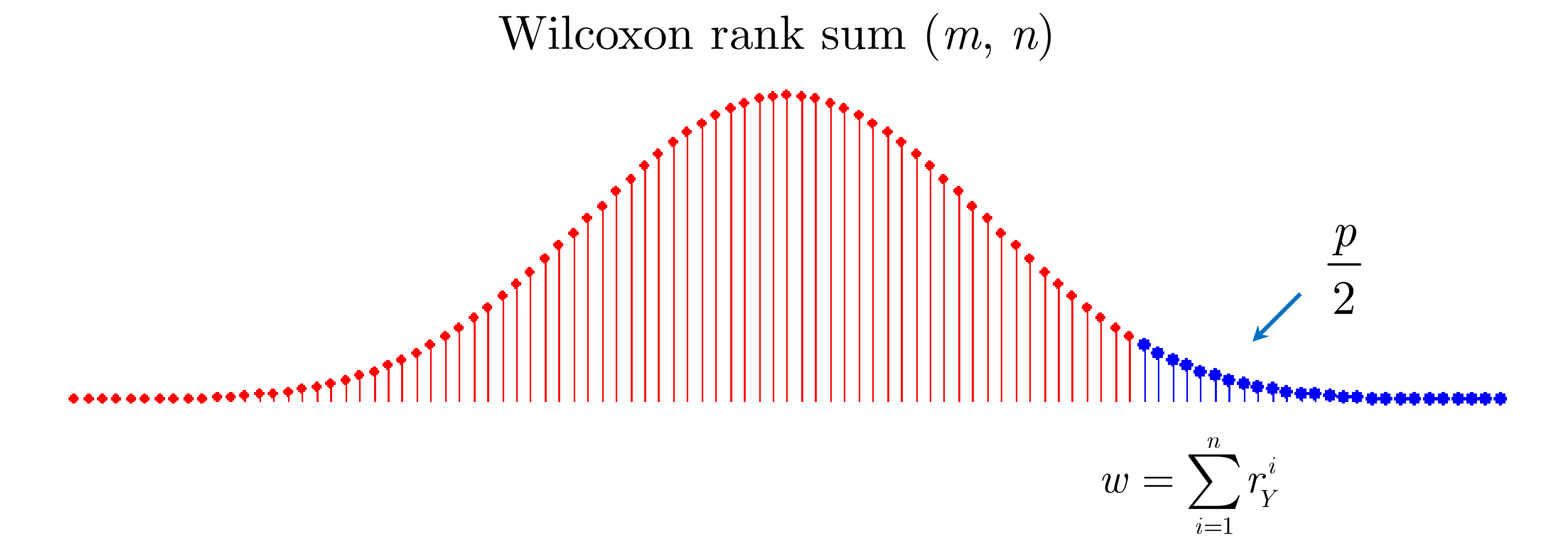
$$W_Y = \sum_{i=1}^{n} R_Y^i$$

p-value:

$$p(\mathbf{x}) = \begin{cases} 2P\left(W_Y \leq \sum_{i=1}^{n} r_y^i\right), & \sum_{i=1}^{n} r_y^i \leq \dfrac{m(m+n+1)}{2} \\ 2P\left(W_Y \geq \sum_{i=1}^{n} r_y^i\right), & \sum_{i=1}^{n} r_y^i \geq \dfrac{m(m+n+1)}{2} \end{cases}$$



Wilcoxon rank sum $(m,\ n)$
$\frac{\alpha}{2}$
$\frac{\alpha}{2}$
$w_{m,n,1-\alpha/2}$
$w_{m,n,\alpha/2}$



Wilcoxon rank sum $(m,\ n)$
$\frac{p}{2}$
$w = \sum_{i=1}^{n} r_Y^i$

	H_0	H_1	Median	Symmetry
One-sample		$m \neq m_0$	Sign test	Wilcoxon Signed rank test
	$m = m_0$	$m > m_0$		
		$m < m_0$		
	$m \leq m_0$	$m > m_0$		
	$m \geq m_0$	$m < m_0$		
Two-sample		$m_X \neq m_Y$	Wilcoxon rank sum test (Mann-Whitney test)	
	$m_X = m_Y$	$m_X > m_Y$		
		$m_X < m_Y$		
	$m_X \leq m_Y$	$m_X > m_Y$		
	$m_X \geq m_Y$	$m_X < m_Y$		
Paired-sample		$m_X \neq m_Y$	Sign test	Paired-sample Wilcoxon signed rank test
	$m_X = m_Y$	$m_X > m_Y$		
		$m_X < m_Y$		
	$m_X \leq m_Y$	$m_X > m_Y$		
	$m_X \geq m_Y$	$m_X < m_Y$		

Active group *vs* Placebo group

1

98	113	103	96	110	96	97	96	94	91
97	113	110							

2

77	88	100	83	108	64	74	87	104	59
91	96	99	108	98	83	83	92	110	73
98	89	88	87	102	89	91	108	119	90
105	72	88	117	78	96	80	97	96	92
85	109	108	87	80	88	85	85	82	98
106	114	90	81	61	84	64	79	77	111
80	65	97	89	107	92	94	112	70	110
76	77	64	80	96	93	94	73	87	111
102	94	78	81	85	75	91	113	87	80
90	77	119	113	96	94	88	103	89	84
71	106	77	86	71	104	108	81	117	99
86	96	91	87	77	100	95	83	93	96
94	93	91	106	101	95	109	127	95	83
78	69	82	89	120	109	85	114	94	92
92	94	100	98	104	109	112	66	83	85
124	95	75	100	95	92	95	113	100	95
87	90	92	101	97	82	82	88	90	103
106	114	93	57	76	77	72	82	101	98
96	114	109	76	99	90	95	64	67	69
80	101	103							

3

109	97	112	115	113	110	105	114	115	90
92	104	109	104	115	90				

4

82	103	116	94	87	93	124	126	131	102
115	103	92	105	105	103	92	103	96	133
99	85	103	109	101	97	130	98	101	87
112	92	96	102	89	108	115	83	116	101
93	96	130	113	135	112	90	92	102	102
97	107	130	121	99	102	103	109	105	77
93	97	96	86	110	107	91	113	133	112
86	77	94	134	108	92	101	104	95	81
112	98	91	90	100	93	69	110	91	92
103	103	85	80	93	100	93	91	96	102
110	124	106	100	133	128	126	92	91	92
78	104	117	133	111	110	116	92	106	110
130	116	110	111	94	100	95	94	95	111
99	110	102	116	99	98	107	67	113	102
125	137	97	102	107	95	125	95	107	131
136	90	113	87	105	134	105	110	132	109
97	100	107	81	90	100	115	75	106	116
100	93	132	105	103	135	79	105	134	87
106	102	122	130	105	142	132	136	132	102
99	106	106	96	102	72				

Practice

Test whether the patients in the same group have lower blood pressure

```
> wilcox.test(active.begin, active.end, "g", paired=T)

         Wilcoxon signed rank test

data:  active.begin and active.end
V = 15548.5, p-value < 2.2e-16
alternative hypothesis: true location shift is greater than 0
...

> wilcox.test(active.begin, placebo.begin, "t")

         Wilcoxon rank sum test

data: active.begin and placebo.begin
W = 20264, p-value = 0.2232
alternative hypothesis: true location shift is not equal to 0
...
```

Compare it with t test!

Goodness-of-fit Test

统计学方法及其应用 > 统计学基础 > 随机变量的函数

"A random variable is a quantity whose values are random and to which a probability distribution is assigned."

Tests of distributions

- Q-Q plots illustrate how good two distributions can fit
- Can we trust our intuition?
- We need a measure to quantify the similarity/difference
- We need to test the Goodness-of-Fit

分布检验

- 检验问题

$$X_1, X_2, \cdots, X_n \sim F(x)$$
$$H_0 : F(x) = F_0(x) \Leftrightarrow H_1 : F(x) \neq F_0(x)$$

- 两种处理方式：
  - 离散：Chi-square检验
  - 连续：KS检验

离散化处理

- 将X的样本空间S划分为K个互不相交的部分 $S_1, \cdots, S_K$ 满足

$$\bigcup_{i=1}^{K} S_i = S, S_i \cap S_j = \emptyset$$

- 令 $p_i = Pr(X \in S_i), i = 1, \cdots, K$机抽取n个观测值，X落在$S_i$的数目设为$n_i$服从多项分布

$$P(N_1 = n_1, \cdots, N_K = n_K) = \frac{n!}{n_1! \cdots n_K!} p_1^{n_1} \cdots, p_K^{n_K}$$

离散后的检验问题

- 观察到落入每个区域S_i的个数为n_i, 如何检验多项分布的参数，即
- 离散检验问题

$$H_0: p_i = \theta_i^0, i = 1, \cdots, K \Leftrightarrow H_1: p_i \neq \theta_i^0 \text{ for at least one i.}$$

- 讨论重点：如何构造统计量，如何确定统计量的分布

简单想法

- 多项分布的极大似然为在相应的频率处达到

$$\hat{\theta}_i = \frac{n_i}{n}, L(\hat{\theta}_1, \cdots, \hat{\theta}_k) = \max L(\theta_1, \cdots, \theta_k)$$

- 似然比统计量

$$T = \frac{L(\theta_1^0, \cdots, \theta_k^0)}{L(\hat{\theta}_1, \cdots, \hat{\theta}_k)} = \prod_{i=1}^{k} \left(\frac{\theta_i^0}{\hat{\theta}_i} \right)^{n_i}$$

$$2 \ln T = 2 \sum_{i=1}^{k} n_i (\ln \theta_i^0 - \ln \hat{\theta}_i)$$

简单想法

- Taylor展开到二阶，

$$\ln\theta_2 - \ln\theta_1 \approx (\theta_2-\theta_1)\frac{1}{\theta_1} + \frac{(\theta_2-\theta_1)^2}{2}(-\frac{1}{\theta_1^2})$$

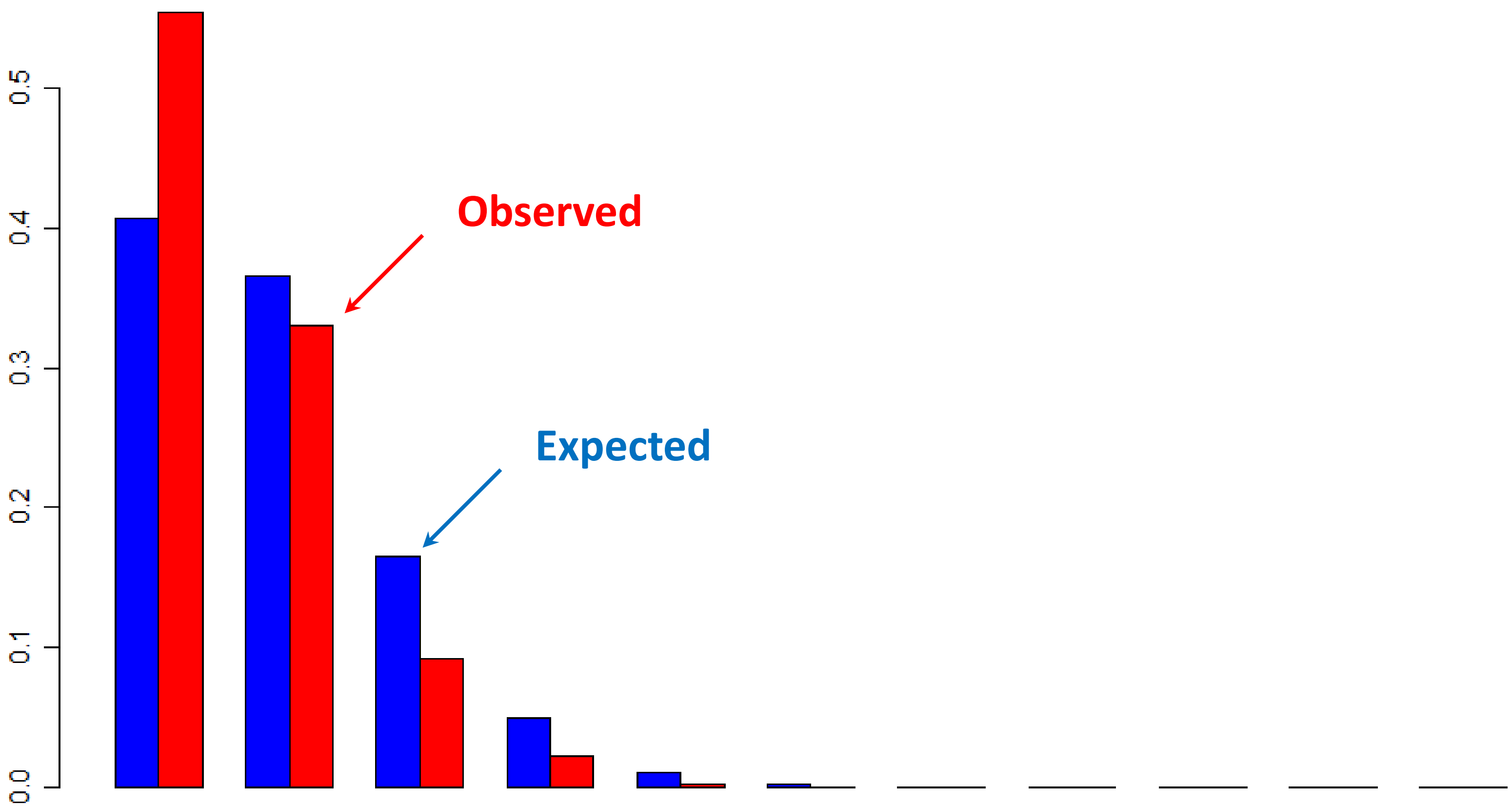
$$\ln\theta_i^0 - \ln\frac{n_i}{n} \approx (\theta_i^0 - \frac{n_i}{n})\frac{1}{\frac{n_i}{n}} + \frac{(\theta_i^0-\frac{n_i}{n})^2}{2}(-\frac{1}{(\frac{n_i}{n})^2})$$

$$= \frac{n\theta_i^0 - n_i}{n_i} - \frac{(n\theta_i^0-n_i)^2}{2n_i^2}$$

- 于是

$$-2\ln T = \sum_i \frac{(n_i - n\theta_i^0)^2}{n_i}$$

Discrete distributions



Chi-square统计量

- 1900年Pearson提出了Chi-square统计量

$$K_n = \sum_{i=1}^{k} \frac{(n_i - n\theta_i^0)^2}{n\theta_i^0}$$

- 和上面的讨论相比，只是把分母的n_i替换成$n\theta_i^0$而这种替换是合理的，因为由大数定律频率是概率的相合估计，即

$$\lim_{n\to+\infty} \left[Pr(\left|\frac{n_i}{n} - \theta_i^0\right| > \epsilon) \right] = 0,\ \forall\epsilon > 0$$

Chi-square统计量的极限分布

- 因为似然比的对数以卡方分布作为极限分布，容易想见Chi-square统计量也以卡方分布作为其极限分布。

- 定理 7.1 在零假设 $H_0 : \hat{\theta}_i = \frac{n_i}{n} = \theta_i^0$，Chi-square统计量有

$$K_n \xrightarrow{L} \chi^2(k-1),\ n \to +\infty$$

自由度修正

- 上面定理中零假设参数是给定的，但更多的时候只是将分布的形式给出，参数需要从数据中估计，这时候需要对自由度进行修正
- 此时问题可以这样表述：理论分布F含有s个位置参数，样本空间分割成k个互不相交的区域，$Pr(X \in S_i) = \pi_i(\theta),\ i = 1, 2, \cdots, k.$
- 根据样本可以得到参数的估计值 $\hat{\theta}_n,$ 由相应的 $\pi_i(\hat{\theta}_n)$ 构造chi-square统计量

自由度修正

$$\hat{K}_n = \sum_{i=1}^{k} \frac{(n_i - \pi_i(\hat{\theta}_n))^2}{n\pi_i(\hat{\theta}_n)}$$

- 定理7.2 (Fisher 1924) 在零假设下，设 $\hat{\theta}_n$ 似然方程组的相合解，则

$$\hat{K}_n \xrightarrow{L} \chi^2_{k-s-1},\ n \to +\infty.$$

Test of the Goodness-of-fit

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} \sim \chi^2_{k-1-j}$$

A level α test would be:

$$\text{Reject } H_0 \text{ if and only if } \chi^2 > \chi^2_{k-1-j,\alpha}$$

The p-value is then

$$p = P(\chi^2_{k-1-j} \geq \chi^2).$$

k is the number of categories, and j is the number of parameters need to be estimated from the data.

An example

Category	i	0	1	2	3	4	5	6
Frequency	o_i	18	34	24	16	3	1	2
Expected	e_i	19.4	31.4	25.4	13.7	5.6	1.8	0.7

Suppose it fits a Poisson distribution

$$P(X = i) = \frac{e^{-\lambda}\lambda^{i}}{i!}$$

$$\hat{\lambda} = \frac{\sum_{i=0}^{6} i o_i}{\sum_{i=0}^{6} o_i} = 1.622$$

$$e_i = P(X = i)\sum_{i=0}^{6} o_i$$

Group cells in which the expected number < 5

$$\chi^2 = \sum_{i=0}^{4} \frac{(o_i - e_i)^2}{e_i} \sim \chi_3^2$$

Original and empirical distributions

The distribution from which a random sample comes from. Let

$$F(x)$$

be its cdf. This distribution is unknown.

Let $X_1, \ldots, X_n$ be a random sample. The empirical cdf $F_n(x)$ is a function of x, defined by

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} I(x_i \leq x),$$

i.e., the fraction of $X_i s$ that are less than or equal to x.

Kolmogorov-Smirnov limit theorem

Let $F(x)$ be the cdf of a distribution. Let $F_n(x)$ be the ecdf of a random sample of size n that comes from $F(x)$. Define

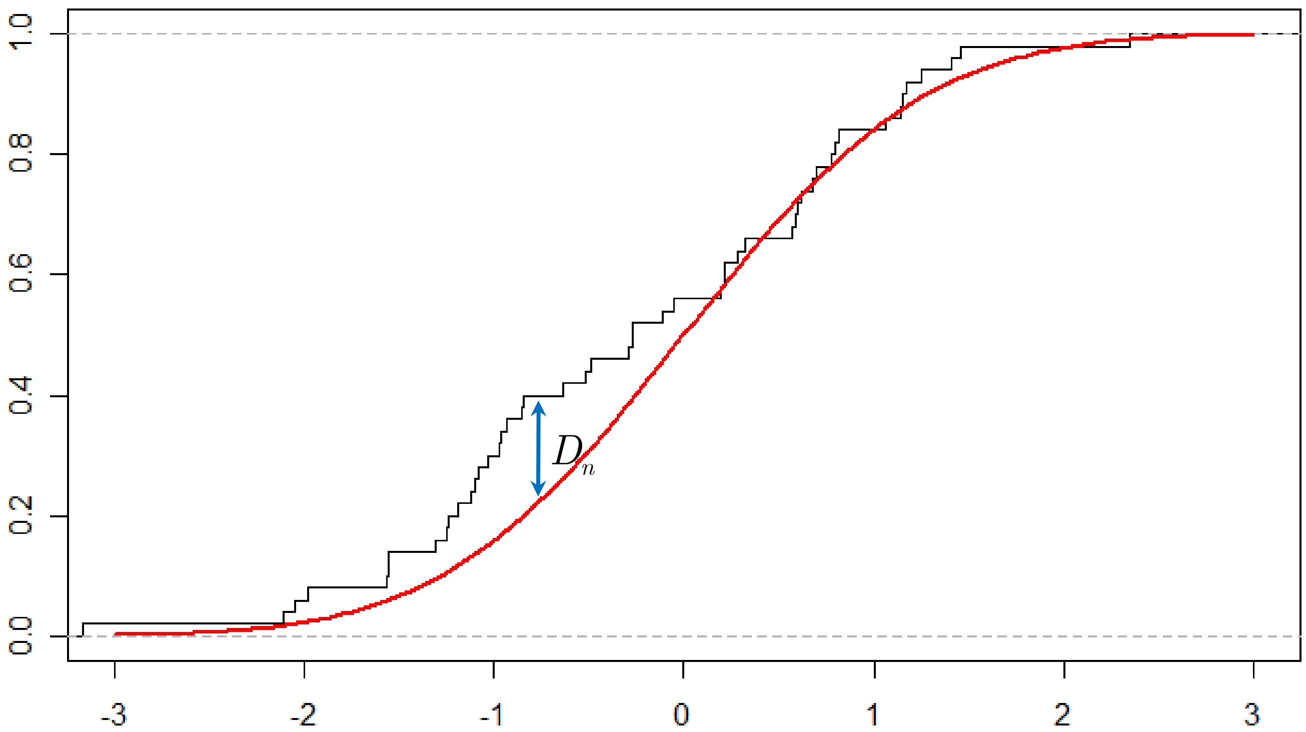
$$D_n = \sup_{x \in \Re} \left| F_n(x) - F(x) \right|.$$

Then

$$\lim_{n \to \infty} P\left(D_n \le \frac{\lambda}{\sqrt{n}} \right) = 1 - 2\sum_{i=1}^{\infty} (-1)^{i+1} e^{-2i^2\lambda^2}, \lambda > 0.$$

D_n is the measure we want to find!

An illustration



Hypothesis testing

Let $F_0(x)$ be the cdf of a distribution. Let $F_n(x)$ be the ecdf of a random sample of size n. We like to test

$$H_0 : F(x) = F_0(x) \quad \text{versus} \quad F(x) \neq F_0(x)$$

Define a test statistic as

$$D = \sup_{x \in \Re} \left| F_n(x) - F_0(x) \right|$$

Under the null, $F(x) = F_0(x)$, the Kolmogorov theorem suggests to reject H_0 when D is large. Detailed calculation of the p-value is complicated.

Practice

```
> x <- rnorm(100)
> ks.test(x, "pnorm")

        One-sample Kolmogorov-Smirnov test

data:  x
D = 0.0981, p-value = 0.2909
alternative hypothesis: two-sided


> x <- rgamma(100, 4, 5)
> ks.test(x, "pgamma", 4, 5)

        One-sample Kolmogorov-Smirnov test

data:  x
D = 0.0686, p-value = 0.7337
alternative hypothesis: two-sided
```

Practice

```
> x <- rnorm(100, 0, 1)
> y <- rnorm(200, 0, 2)
> ks.test(x, y)

          Two-sample Kolmogorov-Smirnov test

data:  x and y
D = 0.215, p-value = 0.004210
alternative hypothesis: two-sided

> x <- rnorm(100, 1, 1)
> y <- rnorm(100, 0, 1)
> ks.test(x, y, alternative="l")

          Two-sample Kolmogorov-Smirnov test

data:  x and y
D^- = 0.37, p-value = 1.134e-06
alternative hypothesis: the CDF of x lies below that of y
```

	H_0	H_1	Normality	Approximation
One-sample	$F = F_0$	$F \neq F_0$	Shapiro test	KS test
		$F > F_0$		
		$F < F_0$		
	$F \leq F_0$	$F > F_0$		
	$F \geq F_0$	$F < F_0$		
Two-sample	$F_X = F_Y$	$F_X \neq F_Y$		KS test
		$F_X > F_Y$		
		$F_X < F_Y$		
	$F_X \leq F_Y$	$F_X > F_Y$		
	$F_X \geq F_Y$	$F_X < F_Y$		

Likelihood Ratio Tests

统计学方法及其应用 统计学基础 随机变量的函数

"A random variable is a quantity whose values are random and to which a probability distribution is assigned."

Asymptotic distribution of the LRT

At the MLE point $\hat{\theta}$,

$$l(\theta \mid \mathbf{x}) = l(\hat{\theta} \mid \mathbf{x}) + l'(\hat{\theta} \mid \mathbf{x})(\theta - \hat{\theta}) + l''(\hat{\theta} \mid \mathbf{x})\frac{(\theta - \hat{\theta})^2}{2!} + \cdots$$

Because $\hat{\theta}$ is MLE, $l'(\hat{\theta} \mid \mathbf{x}) = 0$, then at the true parameter θ_0 (under H_0),

$$-2l(\theta_0 \mid \mathbf{x}) \approx -2l(\hat{\theta} \mid \mathbf{x}) - l''(\hat{\theta} \mid \mathbf{x})(\theta_0 - \hat{\theta})^2$$

Therefore

$$\begin{aligned} -2\log\lambda(\mathbf{x}) &= -2l(\theta_0 \mid \mathbf{x}) + 2l(\hat{\theta} \mid \mathbf{x}) \approx -l''(\hat{\theta} \mid \mathbf{x})(\theta_0 - \hat{\theta})^2 \\ &= \left[-l''(\hat{\theta} \mid \mathbf{x}) / n\right]\left[\sqrt{n}(\hat{\theta} - \theta_0)\right]^2 \\ &= v(\theta_0)\underbrace{\left[-l''(\hat{\theta} \mid \mathbf{x}) / n\right]}_{\to 1/v(\theta_0)}\underbrace{\left[(\hat{\theta} - \theta_0)\sqrt{n} / \sqrt{v(\theta_0)}\right]^2}_{\to \mathrm{N}(0,1)} \\ &\to \chi_1^2 \end{aligned}$$

Asymptotic distribution of the LRT

For hypothesis testing

$$H_0 : \theta = \theta_0 \quad \text{versus} \quad H_1 : \theta \neq \theta_0$$

suppose $X_1, \ldots, X_n$ are iid $f(x \mid \theta)$, $\hat{\theta}$ is the MLE of θ. Then, under the null, as $n \to \infty$,

$$-2 \log \lambda(\mathbf{x}) \to \chi_1^2 \text{ in distribution.}$$

For hypothesis testing

$$H_0 : \theta \in \Theta_0 \quad \text{versus} \quad H_1 : \theta \notin \Theta_0$$

Under the null, as $n \to \infty$,

$$-2 \log \lambda(\mathbf{x}) \to \chi_\nu^2 \text{ in distribution,}$$

where ν is the difference between the number of free parameters specified by $\theta \in \Theta_0$ and that specified by $\theta \in \Theta$.

A simulation study

For random sample from a Bernoulli (p) population, consider the testing

$$H_0 : p = p_0 = 0.5 \quad \text{versus} \quad H_1 : p \neq p_0$$

Under the null,

$$l(p_0 \mid \mathbf{x}) = n_1 \log p_0 + (n - n_1)\log(1 - p_0)$$

Under the whole parameter space

$$l(\hat{p} \mid \mathbf{x}) = n_1 \log(n_1 / n) + (n - n_1)\log((n - n_1) / n)$$

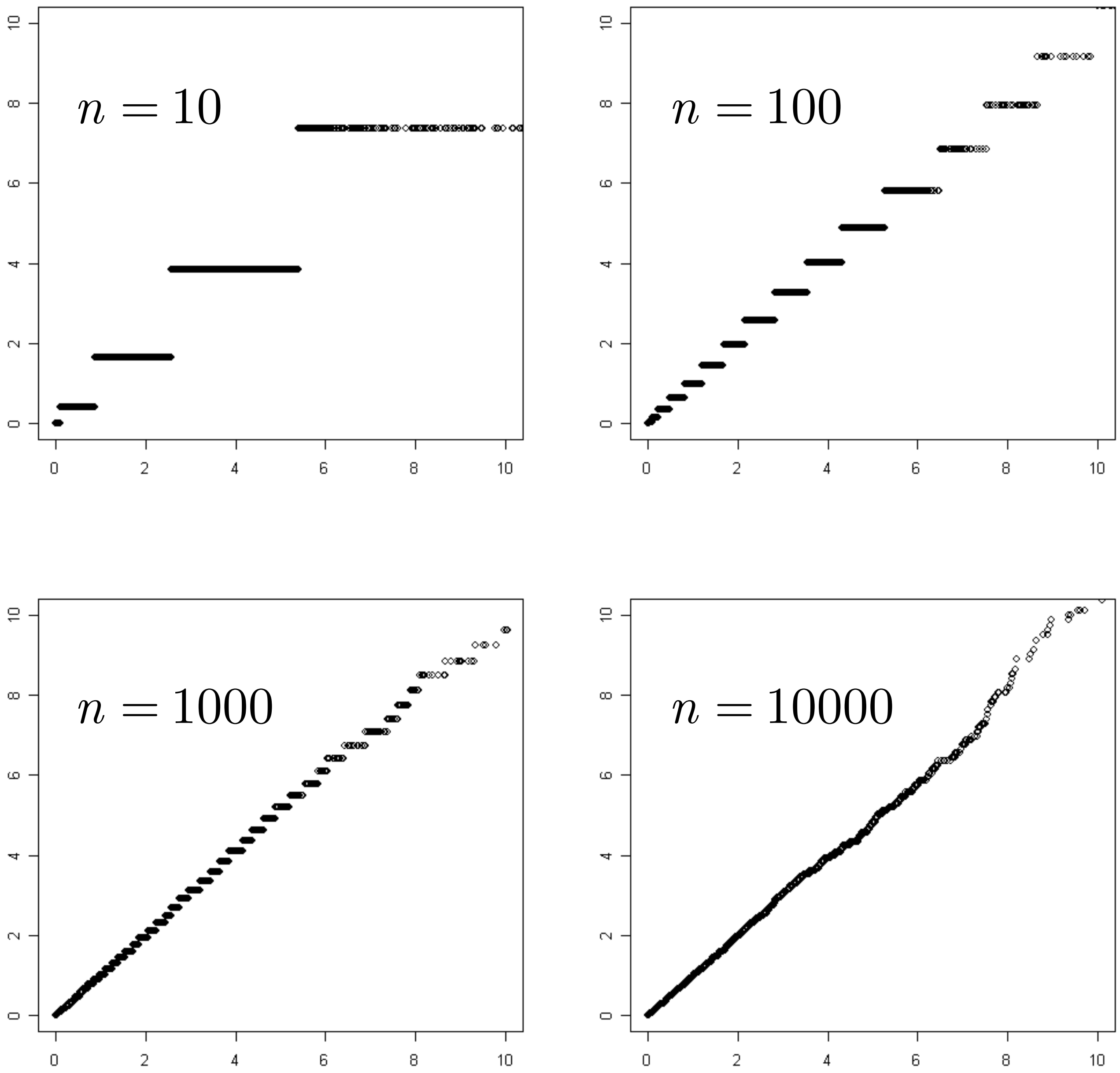
Therefore,

$$\begin{aligned} -2\log\lambda(\mathbf{x}) &= 2n_1 \log\left(\frac{n_1}{np_0}\right) + 2(n - n_1)\log\left(\frac{n - n_1}{n(1 - p_0)}\right) \\ &= 2n_1 \log n_1 + 2(n - n_1)\log(n - n_1) - 2n\log n + 2n\log 2 \end{aligned}$$

when $n \to \infty$,

$$-2\log\lambda(\mathbf{x}) \to \chi_1^2.$$

A simulation study that plots the Q-Q plot for $n = 10, 100, 1000,$ and 10000 is shown in the next slide.


$n = 10$
$n = 100$
$n = 1000$
$n = 10000$

An application

For a sample comes from a multinomial trial distribution that have m cells.

Let $\boldsymbol{\theta} = (\theta_j),\ j = 1,\ldots,m$ be the cell probabilities. So $\sum_{j=1}^{m} \theta_j = 1$.

$$H_0 : \theta_1 = \theta_2 = \cdots = \theta_m \quad \text{versus} \quad H_1 : H_0 \text{ is not true}$$

Under the null,

$$l(\boldsymbol{\theta}_0 \mid \mathbf{x}) = \sum_{j=1}^{m} n_j \log \theta_j = \sum_{j=1}^{m} n_j \log(1/m) = -n \log m$$

Under the whole parameter space

$$l(\hat{\boldsymbol{\theta}} \mid \mathbf{x}) = \sum_{j=1}^{m} n_j \log \hat{\theta}_j = \sum_{j=1}^{m} n_j \log(n_j / n) = \sum_{j=1}^{m} n_j \log n_j - n \log n$$

Therefore,

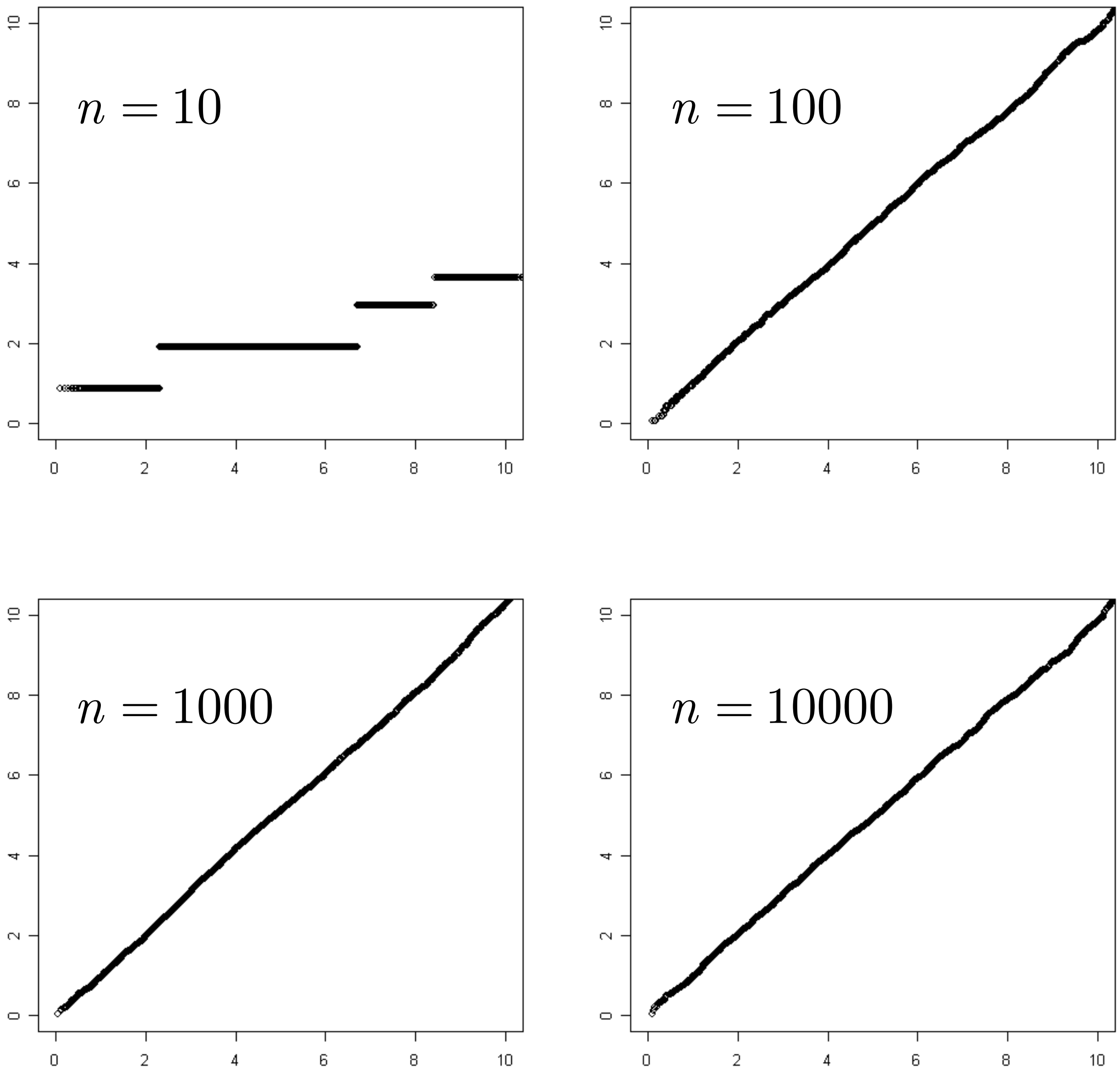
$$-2 \log \lambda(\mathbf{x}) = 2n \log(m/n) + 2\sum_{j=1}^{m} n_j \log n_j,$$

when $n \to \infty$,

$$-2 \log \lambda(\mathbf{x}) \to \chi^2_{m-1}.$$

A level α test on the basis of this statistic would reject H_0 if

$$-2 \log \lambda(\mathbf{x}) > \chi^2_{m-1,\alpha}$$



$n = 10$
$n = 100$
$n = 1000$
$n = 10000$

Gambling

	1	2	3	4	5	6	χ^2	*p*-value
1	11	17	25	9	28	10	19.5498	**0.0015**
2	9	13	13	11	29	25	19.2459	**0.0017**
3	23	25	6	15	11	20	17.8200	**0.0032**
4	16	11	16	7	28	22	17.3693	**0.0039**
5	21	22	5	20	20	12	16.5844	**0.0054**
6	14	11	8	29	18	20	16.4221	**0.0057**
7	21	8	20	8	19	24	15.9945	**0.0069**
8	27	8	12	24	15	14	15.8836	**0.0072**
9	17	11	15	17	17	23	4.5336	**0.4754**
10	16	15	18	15	14	22	2.4766	**0.7800**
11	16	16	17	17	17	17	0.0805	**0.9999**

	H_0	H_1	Normality	Approximation
One-sample	$F = F_0$	$F \neq F_0$	Shapiro test	KS test
		$F > F_0$		
		$F < F_0$		
	$F \leq F_0$	$F > F_0$		
	$F \geq F_0$	$F < F_0$		
Two-sample	$F_X = F_Y$	$F_X \neq F_Y$		KS test
		$F_X > F_Y$		
		$F_X < F_Y$		
	$F_X \leq F_Y$	$F_X > F_Y$		
	$F_X \geq F_Y$	$F_X < F_Y$		