第七章

假设检验

- 统计推断
  - 估计
    - 点估计
    - 区间估计
  - 假设检验
    - 单侧检验
    - 双侧检验

例、两厂生产同一产品，其重量指标都服从正态分布，按规定其均值应该等于 **120**（g）。从甲厂抽出 **5** 件产品测得：**119**，**120**，**119.2**，**119.7**，**119.6**。从乙厂也抽出 **5** 件产品测得：**110.5**，**106.3**，**122.2**，**113.8**，**117.2**，要问这两家工厂的产品是否都符合国家标准

（**1**）先提出假设：$\mu = 120$

（**2**）在给定的置信水平下，利用抽样数据，

判断μ是否落在置信区间（接受区间）

（**3**）做出判断（接受或拒绝原假设）

7.1 假设检验基本概念与一般步骤

（一）提出假设（**Hypothesis)**

原假设H_0,备选假设H_1

H_0与H_1互不相容,

例 **7.1.1** 某药厂包装硼酸粉，规定每袋净重为**0.5**（**kg**），设每袋重量服从正态分布，标准差$\sigma = 0.014$ (**kg**)。为检验包装机的工作是否正常，随机抽取 **10** 袋，称得净重分别为：

0.496	**0.510**	**0.515**	**0.506**	**0.518**
0.512	**0.524**	**0.497**	**0.488**	**0.511**

问这台包装机的工作是否正常

设每袋净重为随机变量 X，则 $X \sim N(\mu, 0.014^2)$

$H_0: \mu = \mu_0$，$H_1: \mu \neq \mu_0$

双侧检验

例 7.1.2　有一厂商声称，有 **75%**以上的用户对其产品质量感到满意。在对 **60** 名用户的调查中，有 **50** 人对该厂产品质量表示满意，问调查数据是否充分支持该厂商的说法?

设用户满意率为 p

H_0: $p \le 0.75$　　H_1: $p > 0.75$

单侧检验

原假设遵循 3 个原则：

（1）等号原则

（2）尊重原假设原则

（3）控制严重后果原则

（二）构造统计量U，由样本观测值计算对应的统计量测试值，并做出判断

例 7.1.1 总体$X \sim N(\mu, \sigma_0^2)$，$\sigma_0^2 = 0.014^2$，

H_0: $\mu = \mu_0 = 0.5$， H_1: $\mu \neq \mu_0$

给定置信度为$\alpha = 0.05$

分析： 若H_0真，即$\mu = \mu_0$

对给定的置信度α，选定常数c，

使得$P\{ \ |\bar{X} - \mu_0| \leq c \ \} = 1 - \alpha$

(1)选统计量 $U=\dfrac{\overline{X}-\mu_0}{\sigma_0/\sqrt{n}}$ 和临界值$U_{1-\alpha/2}$，则有

$$P\{|U|\le U_{1-\alpha/2}\}=1-\alpha$$

计算$\hat{U}=\dfrac{\overline{X}-\mu_0}{\sigma_0/\sqrt{n}}$

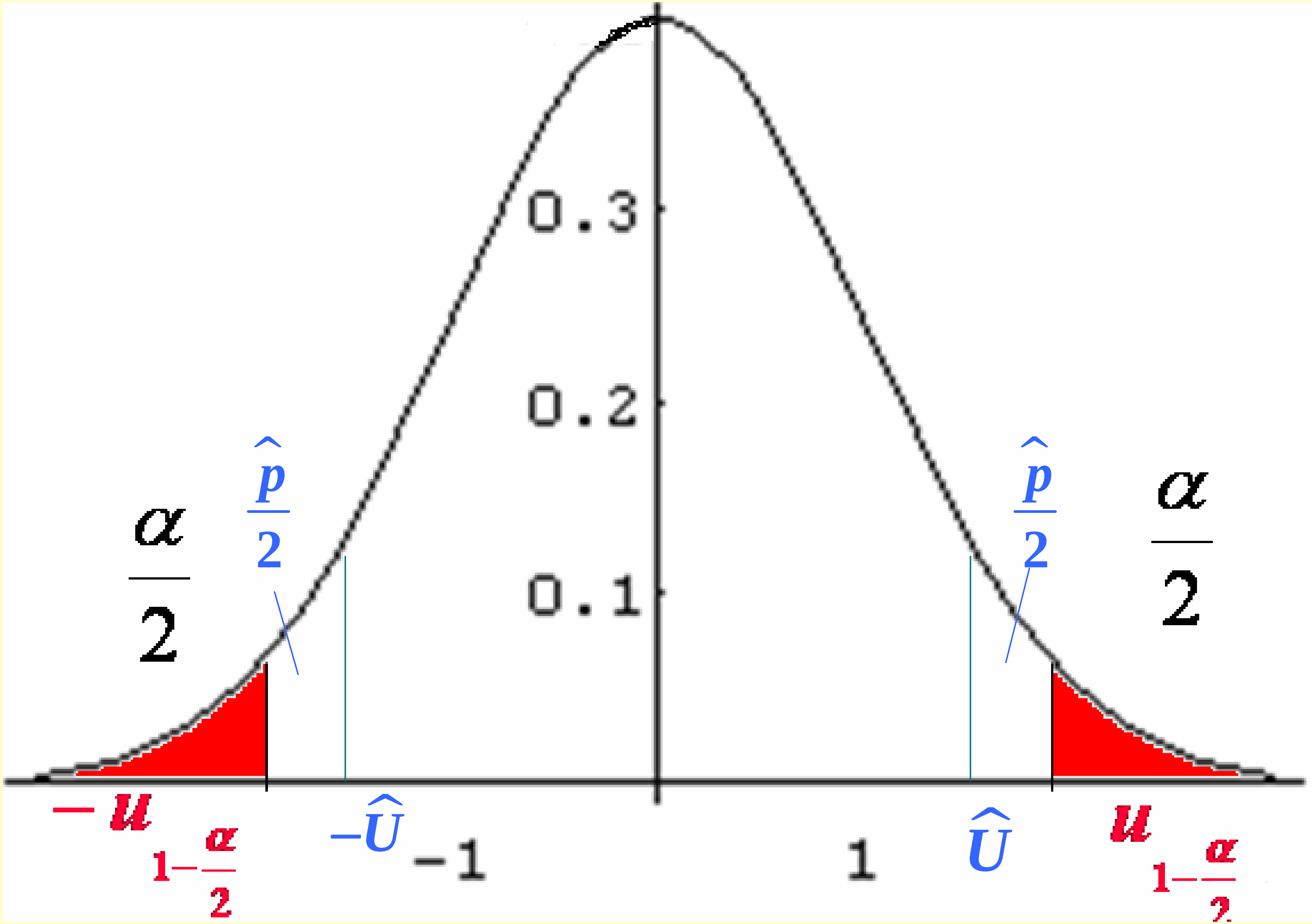
(2)得到区间$W_0:[-U_{1-\alpha/2},U_{1-\alpha/2}]$,

$$W_1:(-\infty,-U_{1-\alpha/2})\cup(U_{1-\alpha/2},+\infty)$$

（3）做出判断：若$\hat{U}\in W_0$,接受；若$\hat{U}\notin W_0$，拒绝

判断的依据：小概率反例否定法

判断的另一种方法：p值检验法


0.3
0.2
0.1
-1
1
$\frac{\alpha}{2}$
$\frac{\alpha}{2}$
$\frac{\hat{p}}{2}$
$\frac{\hat{p}}{2}$
$-\hat{U}$
$\hat{U}$
$-u_{1-\frac{\alpha}{2}}$
$u_{1-\frac{\alpha}{2}}$

假设检验中的两类错误

第一类错误（弃真）： H_0 正确， $\hat{U} \in W_1$

拒真概率（厂方风险） $\alpha = P\{拒绝H_0 \mid H_0真\}$

H_0

第二类错误（取伪）： H_0 错误， $\hat{U} \in W_0$

受伪概率（使用方风险） $\beta = P\{接受H_0 \mid H_0不真\}$

理想方法： α与β 都尽可能地小 不可能

降低错误的方法：

（1）取定α，增大样本容量使β减小

不足：带来检验和试验成本的增加

（2）n给定，规定α，使β尽可能地小

——最优势检验

不足：检测方法不一定存在

（3）限定α，找出$W_0(W_1)$

——显著性检验

不足:只控制α, 不控制β, 使得H_0与H_1地位不平等

控制严重后果原则

$$P\{\hat{U} \in W_1 \mid H_0\text{为真}\} \le \alpha$$

(1) $\hat{U} \in W_1$

（a） H_0错，则推断正确

（b） H_0正确，则推断不正确

$P\{\hat{U} \in W_1 \mid H_0\text{为真}\} \le \alpha$,弃真概率小, 且可控

(2) $\hat{U} \in W_0$

（a） H_0正确，则推断正确

（b） H_0错，则推断不正确

$P\{\hat{U} \in W_0 \mid H_0\text{不真}\} = \beta$,无法控制

弥补的方法： 接受$H_0 \Rightarrow$ 不拒绝H_0

假设检验的一般步骤是：

（1）提出 H_0, H_1，

（2） 选取统计量U，计算测试值$\hat{U}$；

（3）对于给定的α求出临界值，划分W_0和W_1，

使得： $P\{U \in W_1 \mid H_0\text{真}\} \le \alpha$

(4) 做出判断：若$\hat{U} \in W_1$拒绝H_0；

若$\hat{U} \in W_0$不拒绝H_0

正态总体参数的

假设检验

7.2.1 正态总体均值的检验

一、单个总体$N(\mu, \sigma^2)$均值μ的检验

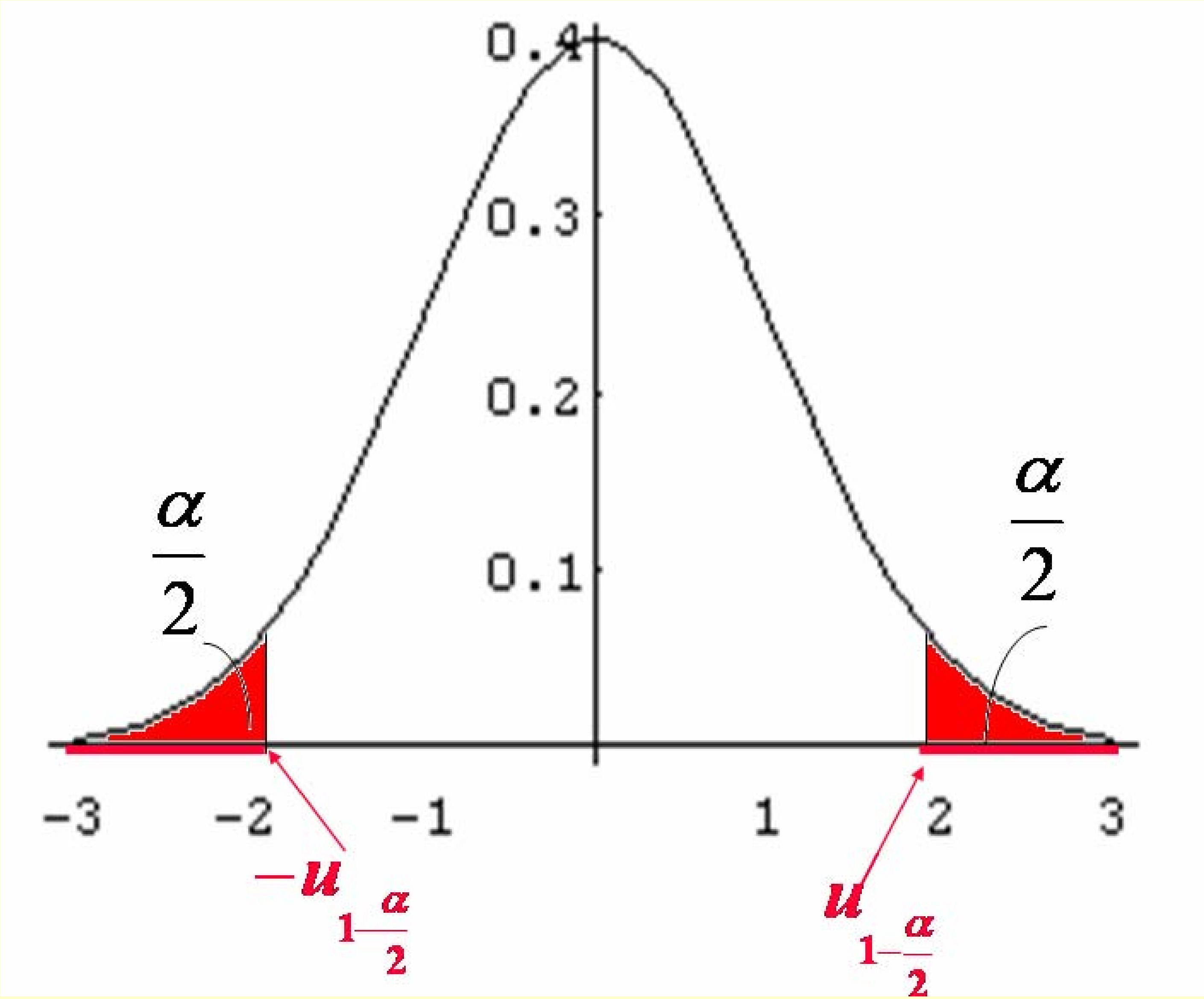
1、$\sigma^2 = \sigma_0^2$

选统计量$U = \dfrac{\overline{X} - \mu_0}{\sigma_0 / \sqrt{n}}$

检验法："U 检验法"或"Z 检验法"

（1）双侧检验：H_0: $\mu = \mu_0$, H_1: $\mu \neq \mu_0$

接受域W_0 : $|U| < U_{1-\alpha/2}$　　拒绝域W_1 : $|U| > U_{1-\alpha/2}$


0.4
0.3
0.2
0.1
$\frac{\alpha}{2}$
$\frac{\alpha}{2}$
-3
-2
-1
1
2
3
$-u_{1-\frac{\alpha}{2}}$
$u_{1-\frac{\alpha}{2}}$

（2）单侧检验

$a)\ H_0: \mu = \mu_0,\ H_1: \mu < \mu_0;$

$b)\ H_0: \mu \geq \mu_0,\ H_1: \mu < \mu_0$

$c)\ H_0: \mu = \mu_0,\ H_1: \mu > \mu_0;$

$d)\ H_0: \mu \leq \mu_0,\ H_1: \mu > \mu_0$

$a)\ H_0: \mu = \mu_0,\ H_1: \mu < \mu_0;$

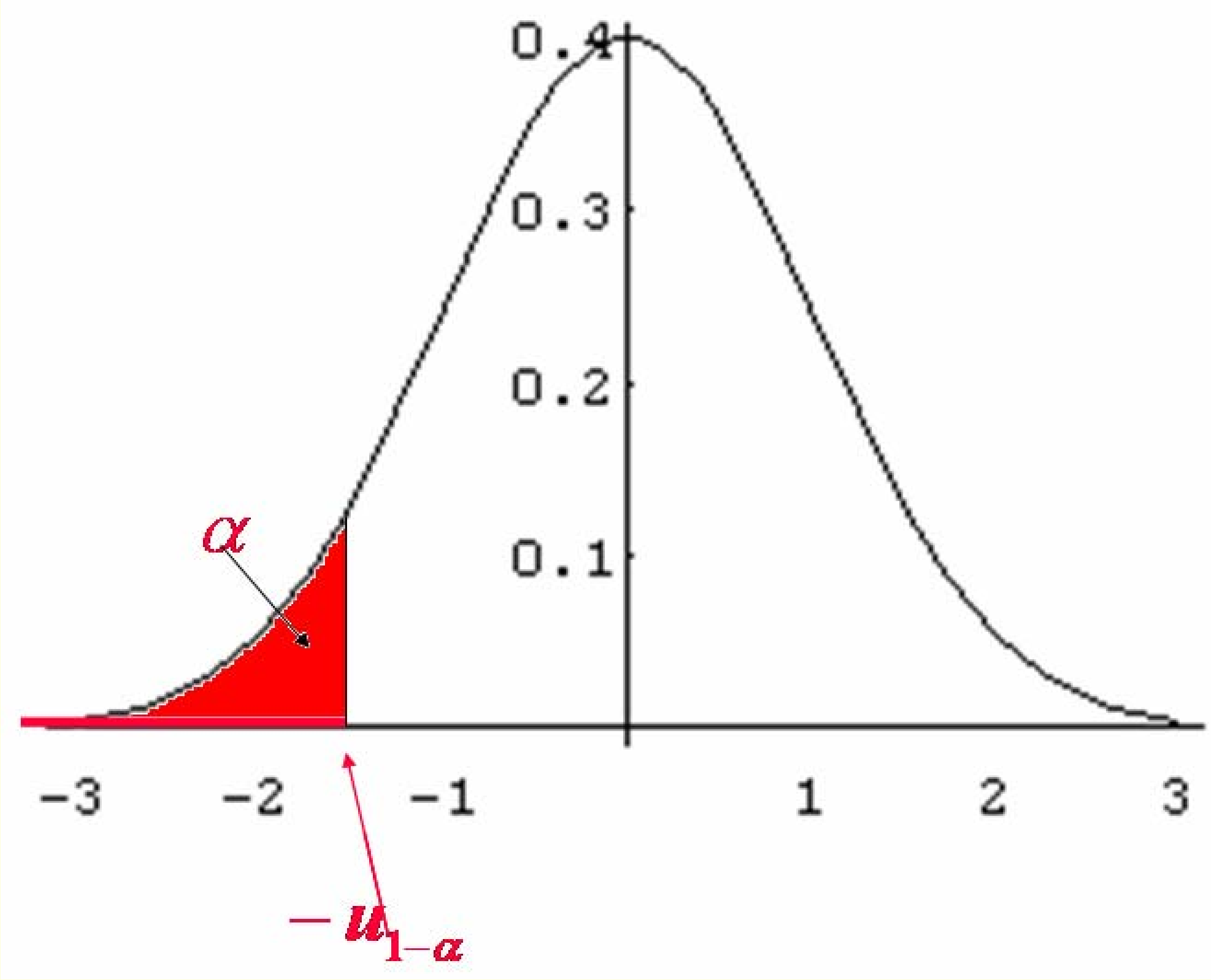
$b)\ H_0: \mu \geq \mu_0,\ H_1: \mu < \mu_0$

$a), b)$ 为左单侧检验

统计量$U = \dfrac{\bar{x} - \mu_0}{\sigma_0 / \sqrt{n}} \overset{H_0\text{成立}}{\sim} N(0,1)$

临界值为$-u_{1-\alpha} = u_\alpha$

拒绝域$W_1 = (-\infty, -U_{1-\alpha})$，接受域$W_0 = [-U_{1-\alpha}, +\infty)$



0.4
0.3
0.2
0.1
α
-3
-2
-1
1
2
3
$-u_{1-\alpha}$

$c)\ H_0: \mu = \mu_0,\ H_1: \mu > \mu_0;$

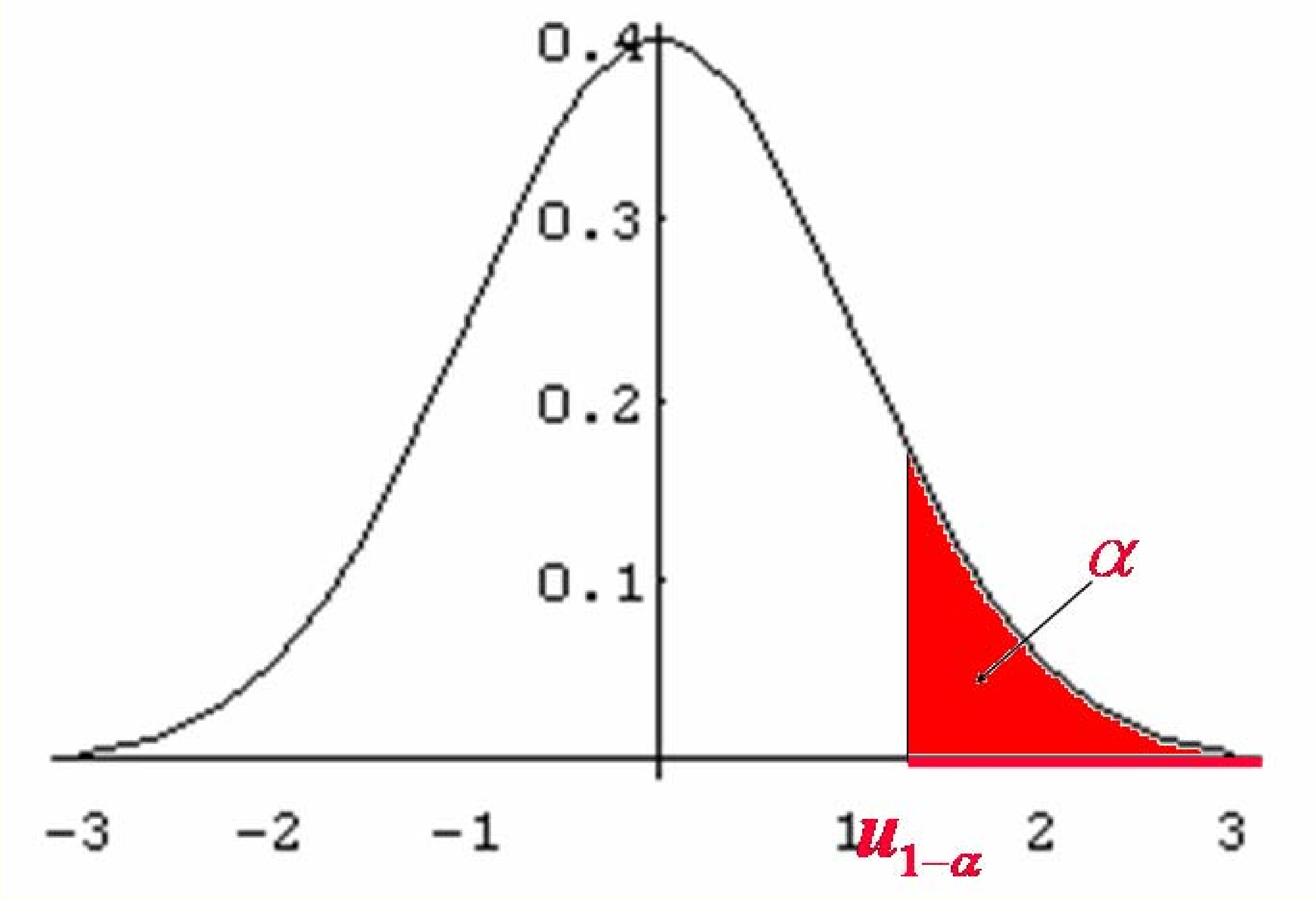
$d)\ H_0: \mu \le \mu_0,\ H_1: \mu > \mu_0$

$c),d)$ 为右单侧检验

$$\text{统计量} U = \frac{\bar{x} - \mu_0}{\sigma_0 / \sqrt{n}} \overset{H_0\text{成立}}{\sim} N(0,1)$$

临界值为 $-u_{1-\alpha} = u_\alpha$

拒绝域 $W_1 = (U_{1-\alpha},\ +\infty)$， 接受域 $W_0 = (-\infty,\ U_{1-\alpha}]$


0.4
0.3
0.2
0.1
α
-3
-2
-1
1
$u_{1-\alpha}$
2
3

例 **7.1.1**　某药厂包装硼酸粉，规定每袋净重为 **0.5**（**kg**），设每袋重量服从正态分布，标准差 $\sigma=0.014$ (**kg**)。为检验包装机的工作是否正常，随机抽取 **10** 袋，称得净重分别为：

0.496	**0.510**	**0.515**	**0.506**	**0.518**
0.512	**0.524**	**0.497**	**0.488**	**0.511**

问这台包装机的工作是否正常

解：设每袋净重为随机变量 X，则 $X \sim N(\mu, 0.014^2)$

$H_0: \mu = \mu_0$，$H_1: \mu \neq \mu_0$

取统计量 $U = \dfrac{\overline{X} - \mu_0}{\sigma_0 / \sqrt{n}} = \dfrac{\overline{X} - 0.5}{0.014 / \sqrt{10}}$

计算得 $\hat{U} = \dfrac{0.5077 - 0.5}{0.014 / \sqrt{10}} = 1.7393$

查表得 $U_{1-\alpha/2} = U_{0.975} = 1.96$

拒绝域为 $W_1 = (-\infty, -1.96] \cup [1.96, +\infty)$

$\because \hat{U} \notin W_1 \quad \therefore$ 不拒绝 H_0

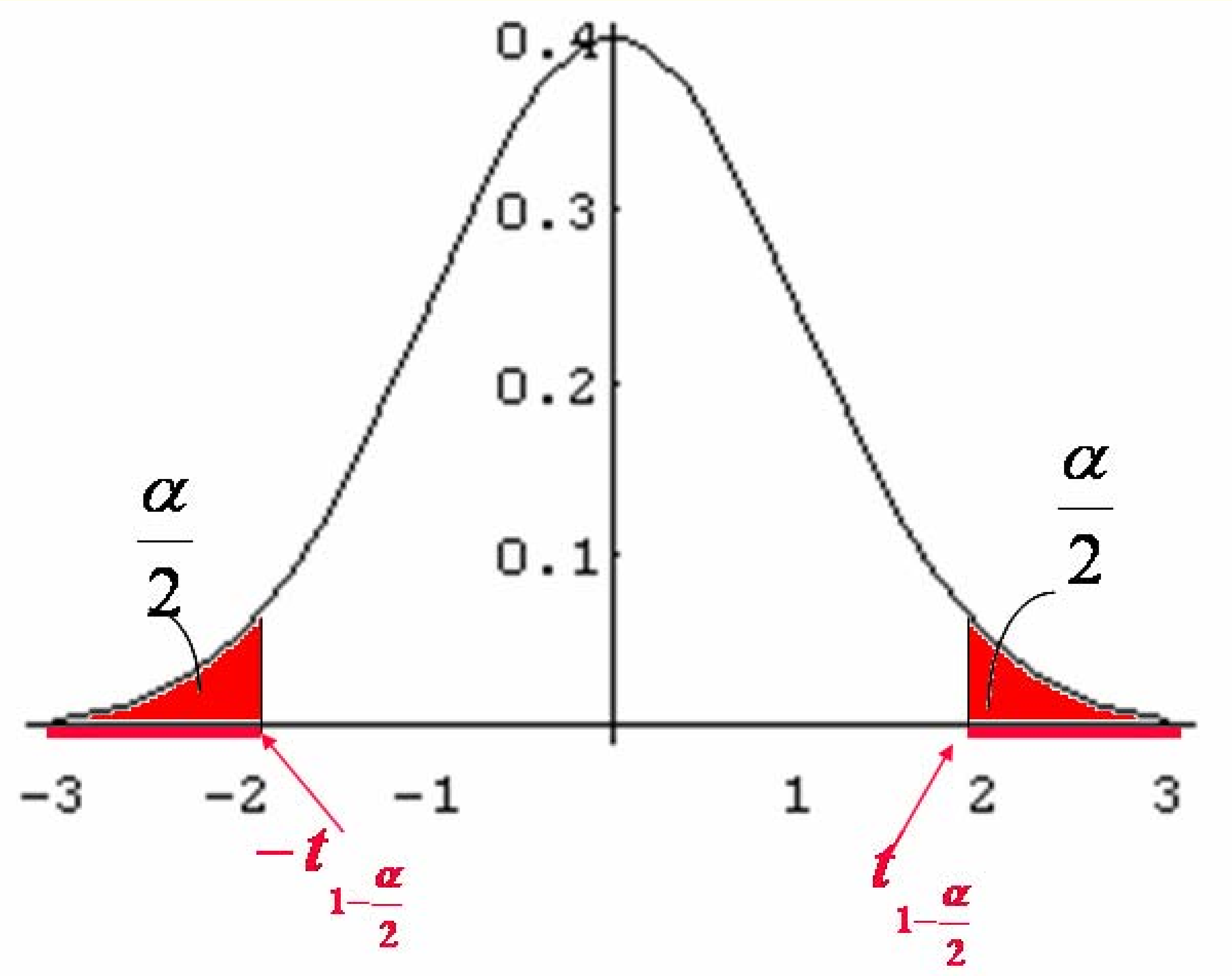
2、σ^2未知

选统计量$T=\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}}$

检验法："t检验法"

（1）双侧检验：H_0: $\mu=\mu_0$, H_1: $\mu\neq\mu_0$

接受域$W_0: |T|<T_{1-\alpha/2}$　　拒绝域$W_1: |T|>T_{1-\alpha/2}$


0.4
0.3
0.2
0.1
$\frac{\alpha}{2}$
$\frac{\alpha}{2}$
−3
−2
−1
1
2
3
$-t_{1-\frac{\alpha}{2}}$
$t_{1-\frac{\alpha}{2}}$

例 **1**、两厂生产同一产品，其重量指标都服从正态分布，按规定其均值应该等于 **120**（g）。从甲厂抽出 **5** 件产品测得：**119，120，119.2，119.7，119.6**。从乙厂也抽出 **5** 件产品测得：**110.5，106.3，122.2，113.8，117.2**，要问这两家工厂的产品是否都符合国家标准

解： $H_0: \mu = 120, H_1: \mu \neq 120$

$\because \sigma^2$ 未知，故采用单个总体的双侧 t 检验法

甲厂，由 $n = 5$, $\overline{X} = 119.5$, $S_x \triangleq S_{n-1} = 0.4$，得测量值

$$\hat{T}_x = \frac{\overline{X} - \mu_0}{S_x / \sqrt{n}} = \frac{119.5 - 120}{0.4 / \sqrt{5}} = -2.795$$

乙厂由 $n = 5$, $\overline{Y} = 114$, $S_y \triangleq S_{n-1} = 6.105$，得测量值

$$\hat{T}_y = \frac{\overline{Y} - \mu_0}{S_y / \sqrt{n}} = \frac{114 - 120}{6.105 / \sqrt{5}} = -2.198$$

临界值 $t_{1-\alpha/2}(n-1)=t_{0.975}(4)=2.776$，

拒域域为 $W_1=(-\infty,-t_{0.975}(4))\cup(t_{0.975}(4),+\infty)$
$$=(-\infty,-2.776)\cup(2.776,+\infty)$$

$\because \hat{T}_x\in\omega_1,\ \hat{T}_y\notin\omega_1$

故推断甲厂产品的均值不等于 120，而乙厂产品的均值不能讲不等于 120

结论不合理性的原因：

1、样本比较粗糙

2、甲产品较稳定，细微差别容易被检验出来。

例2. 某厂生产合金钢，其抗拉强度 $\xi \sim N(\mu, \sigma^2)$,

现在抽查5件样品，测得抗拉强度为

46.8，45.0，48.3，45.1，44.7，

要检验假设强度是否为48？

解：1）检验假设 $H_0: \mu = \mu_0$; $H_1: \mu \neq \mu_0$

2）统计量 $t = \dfrac{\overline{x} - \mu_0}{s/\sqrt{n}} \overset{H_0成立}{\sim} t(4)$

3）计算观测值

$$\overline{x} = 45.98,\ s = 1.535$$

$$|t| = \frac{|45.98 - 48|}{1.535/\sqrt{5}} = 2.942$$

4）与临界值比较 $|t| = 2.942 > 2.7764 = t_{0.025}(4)$

（统计量的值落在拒绝域内）

结论：拒绝原假设，即认为 $\mu \neq 48$。