

九州大学 工学部地球環境工学科  
船舶海洋システム工学コース

海事統計学 第11回 (担当:木村)

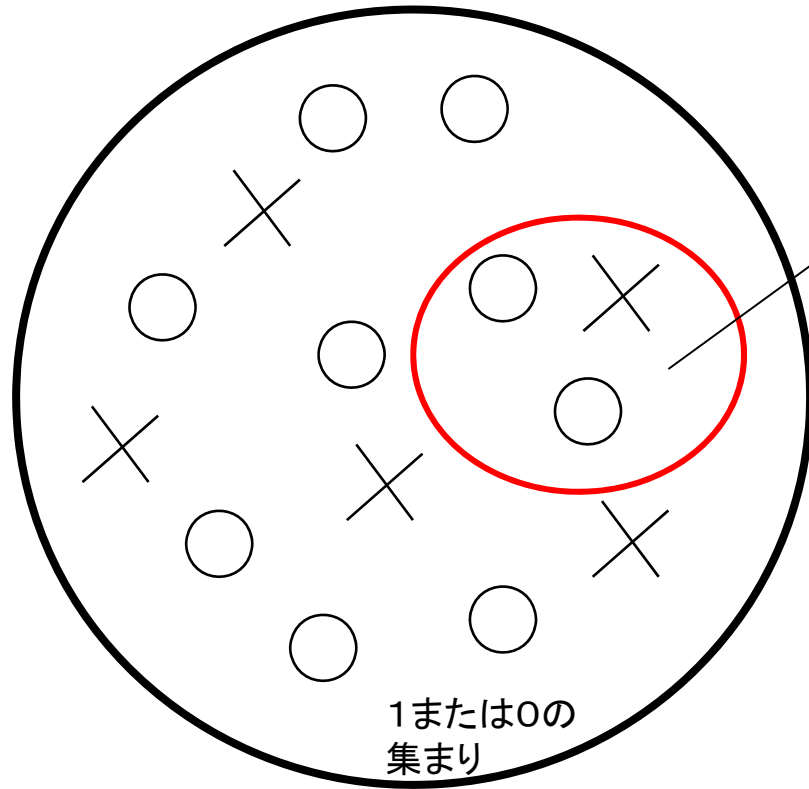
## 仮説検定(1)

授業の資料等は

<http://sysplan.nams.kyushu-u.ac.jp/gen/index.html>

# 【復習】 母比率の大標本

調べる対象となる全体の集団：  
**母集団** (population)



母数(母比率)  $p$   
集団中で1の占める割合

**標本** (sample)

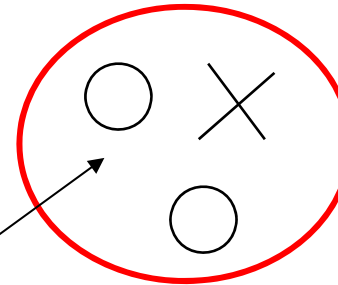
$$x_1, x_2, \dots, x_n$$

確率  $p$  で1をとり、  
それ以外では0を  
とるような確率変数

$x_i$  ベルヌイ分布

期待値  $p$

分散  $p(1-p)$



標本平均(標本比率)

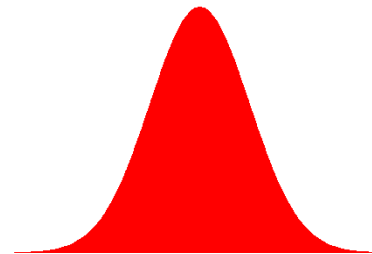
$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

標準化

$$z = \frac{\bar{x} - p}{\sqrt{\frac{p(1-p)}{n}}}$$

の分布は  $N(0, 1)$   
(標準正規分布)

(ただし  $n$  は大きな数)



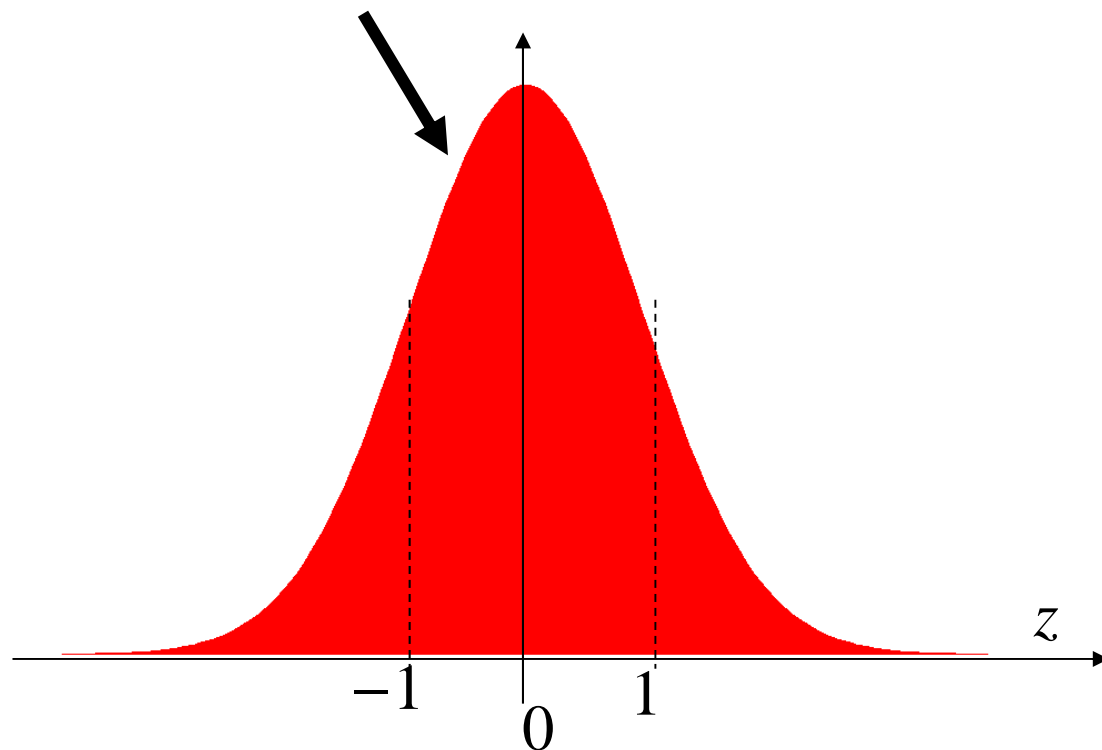
# 【問題】

ある工場で生産される製品の不良率は4%である。  
ある日、製品から2400個を無作為に抽出して検査したところ、  
120個が不良品であった。製造工程に異常が生じたと判断すべきか？

母比率  $p = 0.04$  で、

確率変数  $z = \frac{\bar{x} - p}{\sqrt{\frac{p(1-p)}{n}}}$  は

標準正規分布  $N(0, 1)$  に従うはず



実際に得られた標本平均(標本比率)

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{120}{2400} = 0.05$$

代入

$$z = \frac{\bar{x} - p}{\sqrt{\frac{p(1-p)}{n}}} = \frac{0.05 - 0.04}{\sqrt{\frac{0.04(1-0.04)}{2400}}} = \underline{2.5}$$



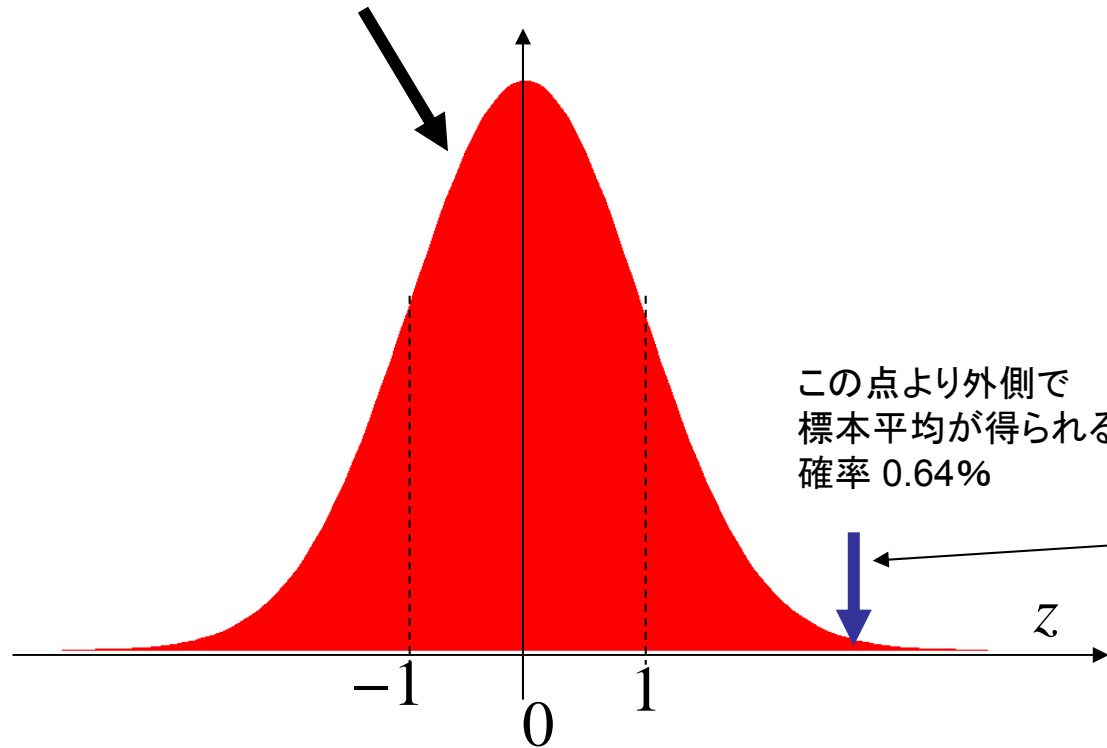
# 【問題】

ある工場で生産される製品の不良率は4%である。  
ある日、製品から2400個を無作為に抽出して検査したところ、  
120個が不良品であった。製造工程に異常が生じたと判断すべきか？

母比率  $p = 0.04$  で、

確率変数  $z = \frac{\bar{x} - p}{\sqrt{\frac{p(1-p)}{n}}}$  は

標準正規分布  $N(0, 1)$  に従うはず



実際に得られた標本平均(標本比率)

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{120}{2400} = 0.05$$

代入

$$z = \frac{\bar{x} - p}{\sqrt{\frac{p(1-p)}{n}}} = \frac{0.05 - 0.04}{\sqrt{\frac{0.04(1-0.04)}{2400}}} = 2.5$$

めったに起きない事象が起こった？

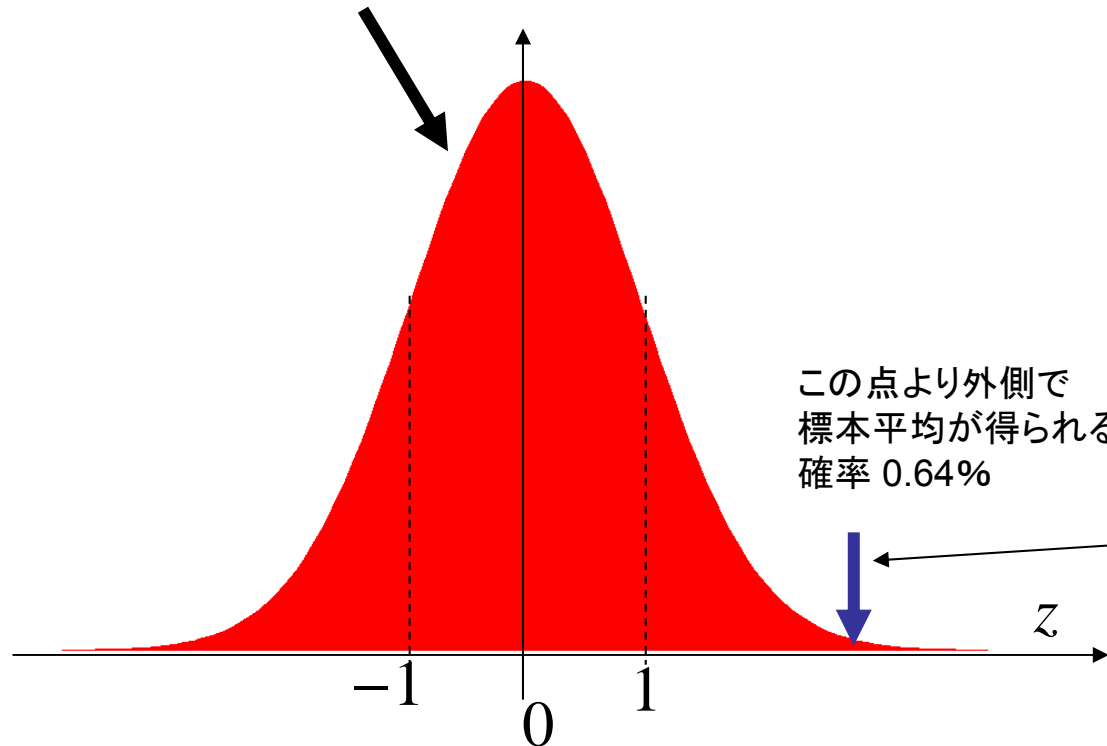
# 【問題】

ある工場で生産される製品の不良率は4%である。  
ある日、製品から2400個を無作為に抽出して検査したところ、  
120個が不良品であった。**製造工程に異常が生じたと判断すべきか？**

母比率  $p = 0.04$  で、

確率変数  $z = \frac{\bar{x} - p}{\sqrt{\frac{p(1-p)}{n}}}$  は

標準正規分布  $N(0, 1)$  に従うはず



実際に得られた標本平均(標本比率)

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{120}{2400} = 0.05$$

代入

$$z = \frac{\bar{x} - p}{\sqrt{\frac{p(1-p)}{n}}} = \frac{0.05 - 0.04}{\sqrt{\frac{0.04(1-0.04)}{2400}}} = 2.5$$

~~めったに起きない事象が起こった？~~

**$p = 0.04$  という仮定が正しくない！  
→ 製造工程に異常と診断**

# 検定とは？

## Hypothesis Testing

標本から得られる情報に基づき、仮説を否定すべきかどうかを判断する統計的方法(仮説検定)

・母集団に関する仮定 = 統計的仮説



否定  
したい仮説



認めたい  
仮説

不良品の例題

異常なし  
母比率  $p = 0.04$

異常あり  
母比率  $p > 0.04$

または  $p \neq 0.04$

・検定により、仮説が正しくないと判断してこれを否定する =



・「めったに起きない事象」として棄却を判断する基準となる確率  $\alpha$  =



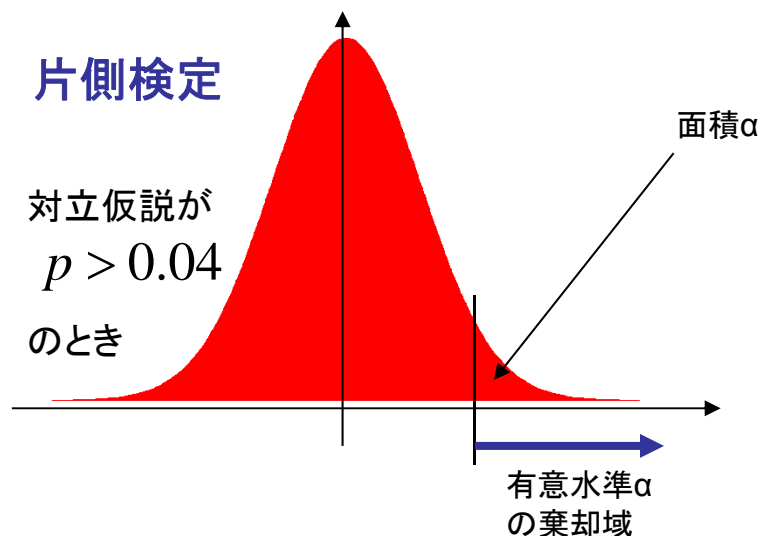
$\alpha$ としては 0.05 や 0.01 をとることが多い

・有意水準に対し、仮説のもとでは実現しにくい標本値の範囲 = 有意水準  $\alpha$  の



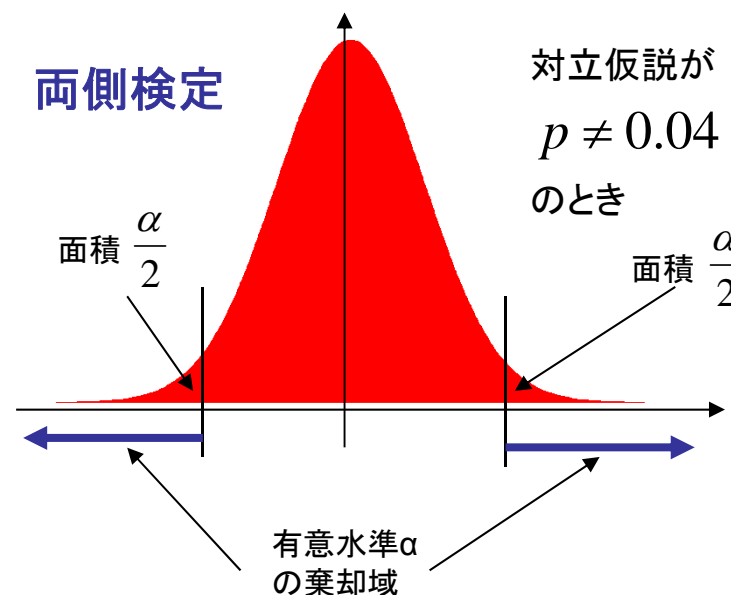
### 片側検定

対立仮説が  
 $p > 0.04$   
のとき



### 両側検定

対立仮説が  
 $p \neq 0.04$   
のとき



# 検定とは？

## Hypothesis Testing

標本から得られる情報に基づき、仮説を否定すべきかどうかを判断する統計的方法(仮説検定)

・母集団に関する仮定 = 統計的仮説

**帰無仮説**

否定  
したい仮説

**対立仮説**

認めたい  
仮説

不良品の例題

異常なし  
母比率  $p = 0.04$

異常あり  
母比率  $p > 0.04$

または  $p \neq 0.04$

・検定により、仮説が正しくないと判断してこれを否定する = **仮説の棄却**

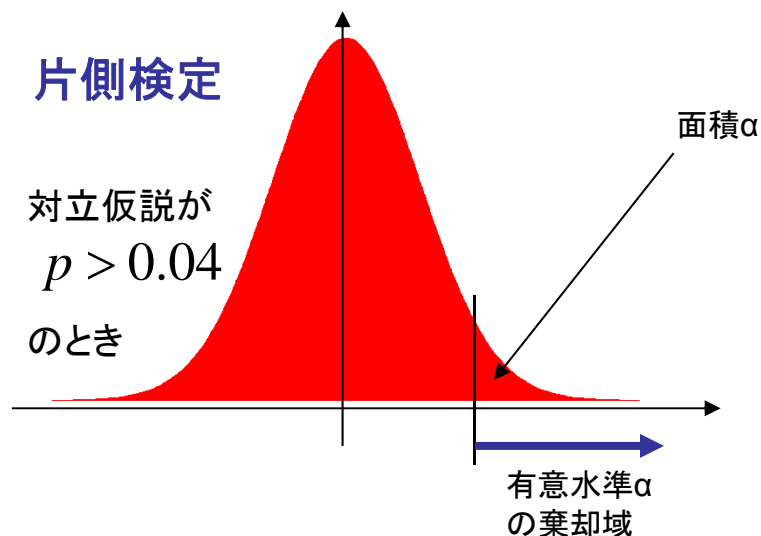
・「めったに起きない事象」として棄却を判断する基準となる確率  $\alpha$  = **有意水準(危険率)**

$\alpha$ としては 0.05 や 0.01 をとることが多い

・有意水準に対し、仮説のもとでは実現しにくい標本値の範囲 = 有意水準  $\alpha$  の **棄却域**

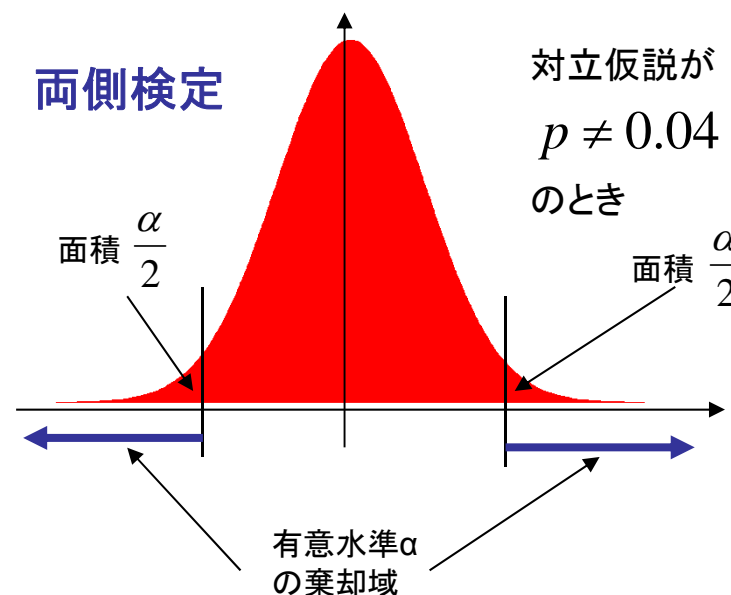
### 片側検定

対立仮説が  
 $p > 0.04$   
のとき



### 両側検定

対立仮説が  
 $p \neq 0.04$   
のとき





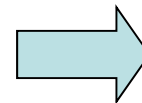
# 検定における「2種類の誤り」

- ・本当は**仮説が正しい**のに、これを**棄却してしまう誤り** = 第1種の誤り  
誤りが生じる確率 =  $\alpha$
- ・本当は**仮説が誤っている**のにこれを**棄却できない誤り** = 第2種の誤り

		真実(母集団の状態)	
		仮説が正しい	仮説が誤り
検定の結論	仮説を否定できない	正	
	仮説を否定		正

# 検定における「非対称性」

仮説が偽であることを立証することはできるが、真であることを積極的に立証することはできない



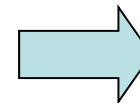
# 検定における「2種類の誤り」

- ・本当は**仮説が正しい**のに、これを**棄却してしまう誤り** = 第1種の誤り  
誤りが生じる確率 =  $\alpha$
- ・本当は**仮説が誤っている**のにこれを**棄却できない誤り** = 第2種の誤り

		真実(母集団の状態)	
		仮説が正しい	仮説が誤り
検定の結論	仮説を否定できない	正	<b>第2種の誤り</b>
	仮説を否定	<b>第1種の誤り</b>	正

# 検定における「非対称性」

仮説が偽であることを立証することはできるが、真であることを積極的に立証することはできない



仮説が棄却されなくても、標本数を増やせば棄却される可能性がある

# 検定によってどんなことが判定できるか？

## ・比率の検定

- (1) 母比率  $P$  が、ある値  $P_0$  に等しいといえるか？
- (2) 比率の差の検定： 2つの異なる母集団の間で、母比率に差があるといえるか？

## ・平均値の検定

- (1) 母集団の平均値  $\mu$  が、ある値  $\mu_0$  に等しいといえるか？
- (2) 平均値の差の検定： 2つの異なる母集団の間で、母平均に差があるといえるか？

## ・分散の検定

- (1) **正規母集団**の分散  $\sigma^2$  が、ある値  $\sigma_0^2$  に等しいといえるか？
- (2) 分散の差の検定： 2つの異なる**正規母集団**の間で、分散に差があるといえるか？

## ・適合度の検定

- (1) 観察されたデータが、特定の分布に一致しているといえるか？

- (2) 2つの母集団の確率分布が異なるものであるかどうか？

分布の種類を問わない  
(ノンパラメトリック)

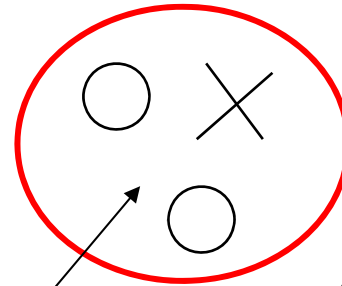
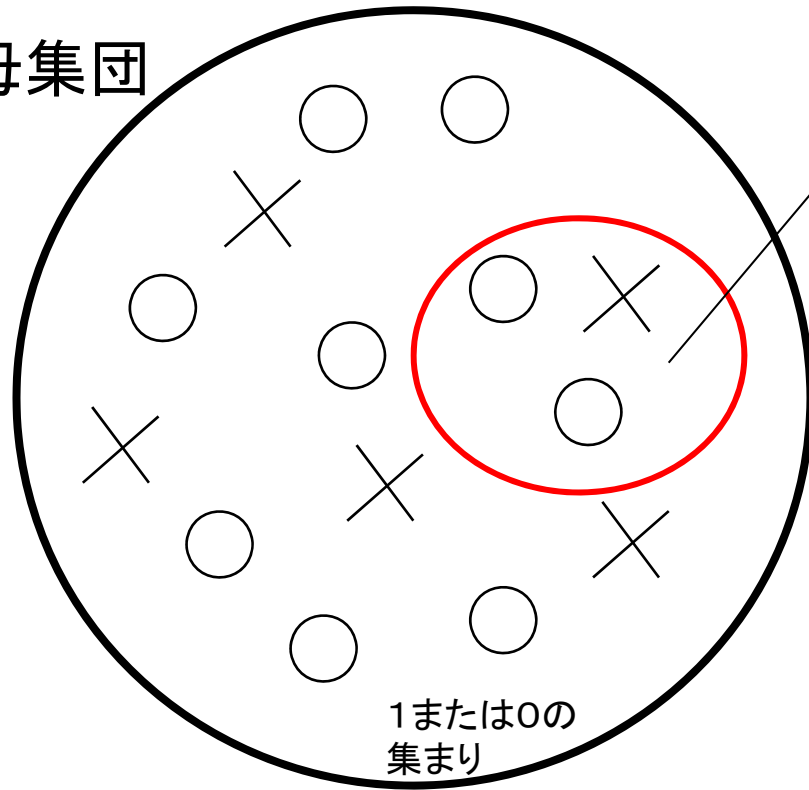
← コルモゴロフ・スミルノフ検定

# 比率の検定

標本  $x_1, x_2, \dots, x_n$

(1) 母比率  $P$  がある値  $p_0$  に等しいといえるかどうか？

母集団



標本平均(標本比率)

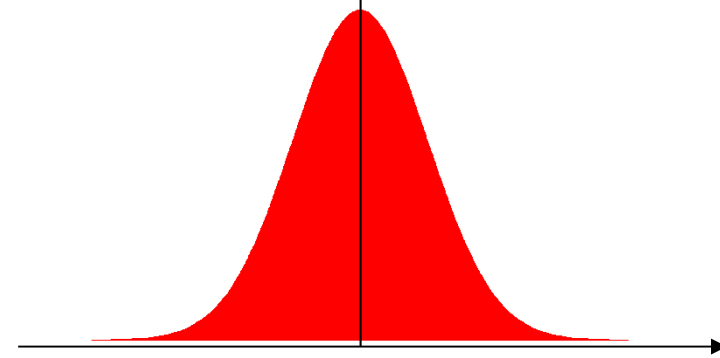
$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

標準化



の分布は  $N(0, 1)$   
(標準正規分布)

(ただし  $n$  は大きな数)



母比率  $p = p_0$  ← 帰無仮説  
集団中で1の占める割合

対立仮説  $p > p_0$  → 片側検定  
(または  $p < p_0$ )

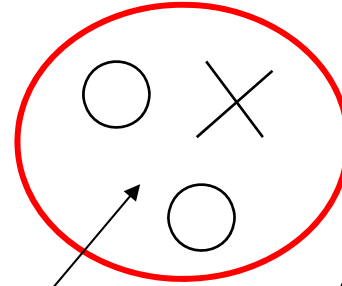
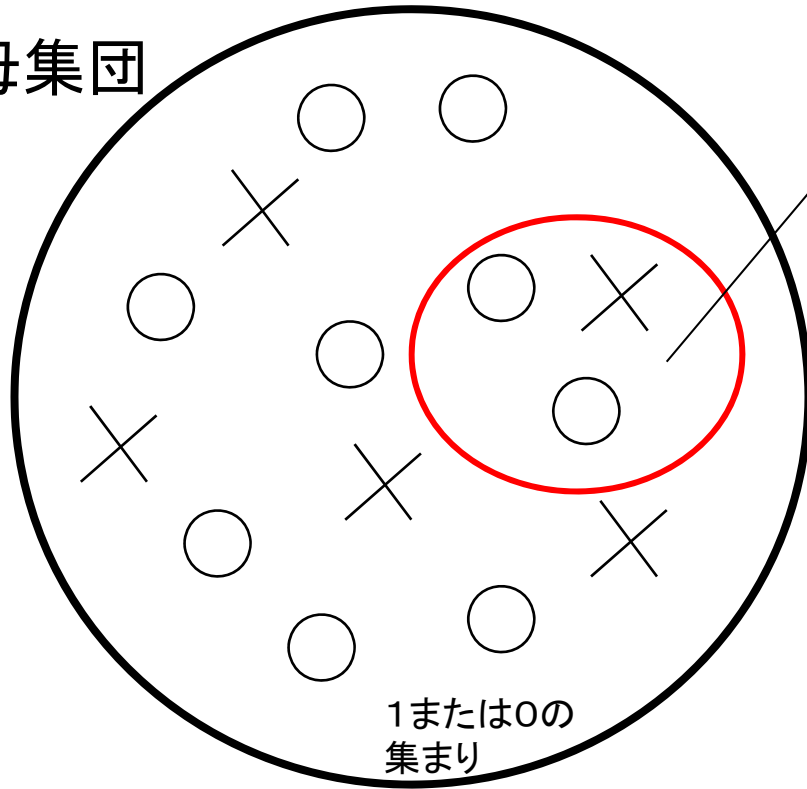
$p \neq p_0$  → 両側検定

# 比率の検定

標本  $x_1, x_2, \dots, x_n$

(1) 母比率  $P$  がある値  $p_0$  に等しいといえるかどうか？

母集団



標本平均 (標本比率)

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

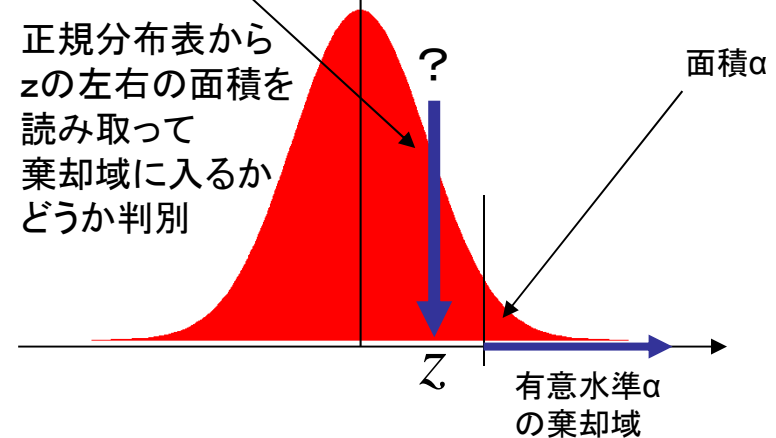
標準化

$$z = \frac{\bar{x} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

の分布は  $N(0, 1)$   
(標準正規分布)

(ただし  $n$  は大きな数)

正規分布表から  $z$  の左右の面積を読み取って棄却域に入るかどうか判別



母比率  $p = p_0$  ← 帰無仮説  
集団中で1の占める割合

対立仮説  $p > p_0$  → 片側検定  
(または  $p < p_0$ )  
 $p \neq p_0$  → 両側検定

有意水準  $\alpha$  を設定し、標本から得られた  $z$  が正規分布の棄却域に入るかどうか判定  
棄却域に入ったら帰無仮説を棄却する

【参考】

サンプル数が少ない場合は  
二項分布を用いて計算



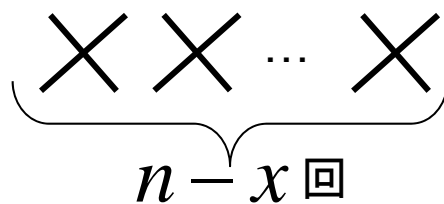
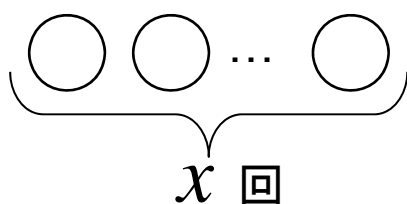
1回の試行で、ある事象の起こる確率が  $p$   
この試行を  $n$  回行う

確率変数  $x$  : 事象の起こった回数  
確率分布  $P(x)$  → 二項分布  $B(n,p)$

(例) 3回サイコロを投げたときの1の目が出る回数 =  $x$

$x$	0	1	2	3
$P(x)$	$\left(\frac{5}{6}\right)^3 = \frac{125}{216}$	$3\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^2 = \frac{75}{216}$	$3\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right) = \frac{15}{216}$	$\left(\frac{1}{6}\right)^3 = \frac{1}{216}$

一般に、確率  $p$  を持つ事象が  $n$  回の試行中  $x$  回起こることを考えると、



このときの確率は  $p^x (1-p)^{n-x}$

$n$  回の試行中に  $x$  回起こるパターンの組合せ (combination) は、

$${}_n C_x = \frac{n!}{x!(n-x)!} \quad \text{よって} \quad P(x) = {}_n C_x p^x (1-p)^{n-x} = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

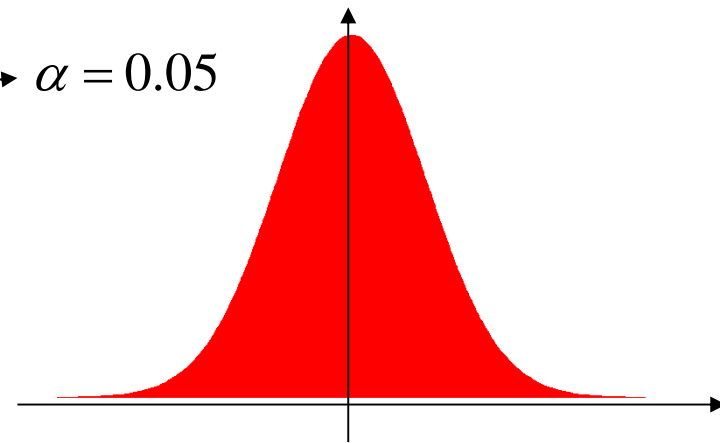
### 【練習問題】

貨物船による輸送においては、船舶の30%において貨物が水による何らかのダメージを受けるといわれる。ランダムに抽出した36隻の貨物船において聞き取りを行ったところ、15隻において貨物に水による何らかのダメージが生じたとの回答が得られた。最初の30%という仮説の数字は修正すべきか？有意水準 5%で検定せよ。

ただし  $\sqrt{21} = 4.58$       $\frac{15}{36} = 0.42$      として計算せよ。

**帰無仮説:**  $p_0 = 0.3$

**対立仮説:**  $p_0 \neq 0.3$







### 【練習問題】

貨物船による輸送においては、船舶の30%において貨物が水による何らかのダメージを受けるといわれる。ランダムに抽出した36隻の貨物船において聞き取りを行ったところ、15隻において貨物に水による何らかのダメージが生じたとの回答が得られた。最初の30%という仮説の数字は修正すべきか？有意水準 5%で検定せよ。

ただし  $\sqrt{21} = 4.58$      $\frac{15}{36} = 0.42$     として計算せよ。

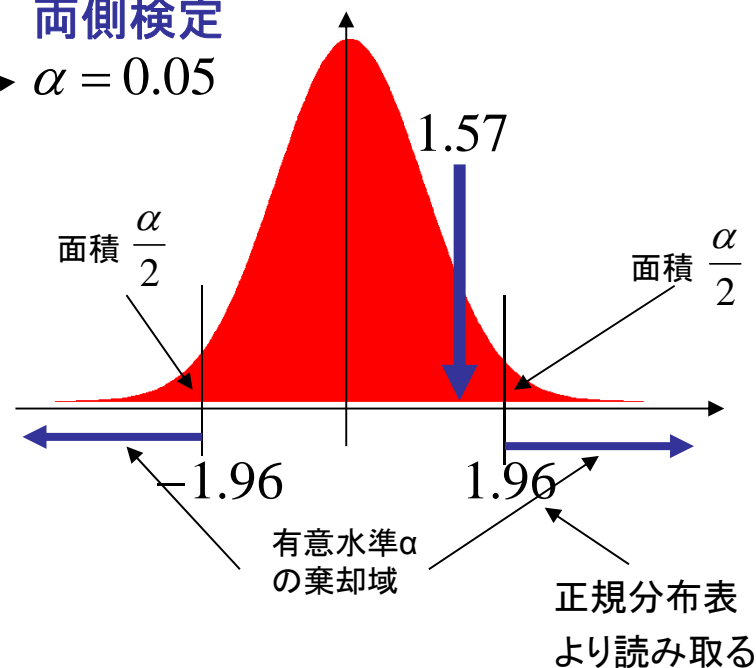
帰無仮説:  $p_0 = 0.3$

対立仮説:  $p_0 \neq 0.3$

両側検定

$\alpha = 0.05$

$$z = \frac{\bar{x} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{\frac{15}{36} - 0.3}{\sqrt{\frac{0.3 \times (1-0.3)}{36}}} \approx \frac{0.42 - 0.3}{\sqrt{\frac{3 \times 7}{6 \times 6 \times 10 \times 10}}} \approx 1.57$$



## 【練習問題】

貨物船による輸送においては、船舶の30%において貨物が水による何らかのダメージを受けるといわれる。ランダムに抽出した36隻の貨物船において聞き取りを行ったところ、15隻において貨物に水による何らかのダメージが生じたとの回答が得られた。最初の30%という仮説の数字は修正すべきか？有意水準 5%で検定せよ。

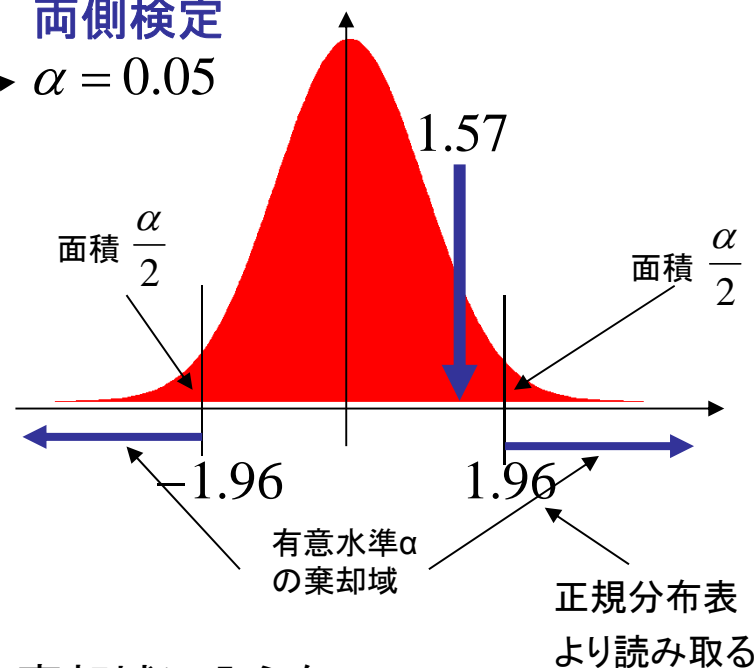
ただし  $\sqrt{21} = 4.58$      $\frac{15}{36} = 0.42$     として計算せよ。

帰無仮説:  $p_0 = 0.3$

対立仮説:  $p_0 \neq 0.3$

$$z = \frac{\bar{x} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{\frac{15}{36} - 0.3}{\sqrt{\frac{0.3 \times (1-0.3)}{36}}} \approx \frac{0.42 - 0.3}{\sqrt{\frac{3 \times 7}{6 \times 6 \times 10 \times 10}}} \approx 1.57$$

両側検定  
 $\alpha = 0.05$

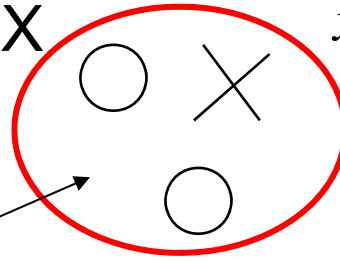


棄却域に入らないので、  
仮説は棄却できない  
＝仮説は修正すべきでない

# 比率の検定

(2) 異なる2つの母集団の間で、母比率に差があるといえるかどうか？

標本X



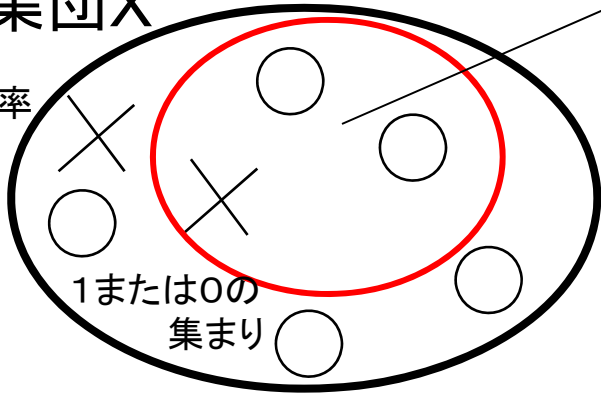
$x_1, x_2, \dots, x_m$

標本平均 (標本比率)

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_m}{m}$$

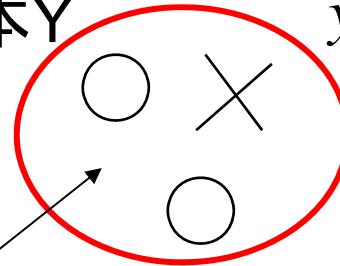
母集団X

母比率  $p_x$



1または0の集まり

標本Y



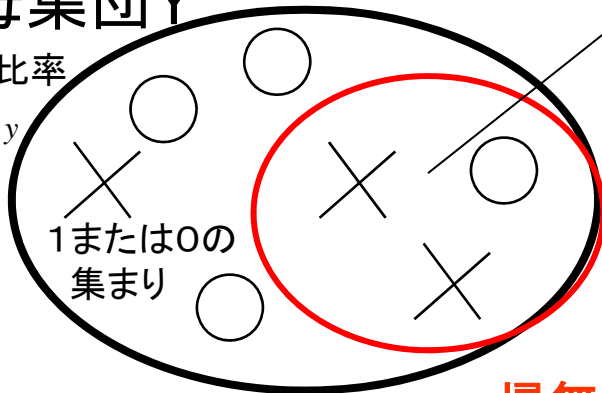
$y_1, y_2, \dots, y_n$

標本平均 (標本比率)

$$\bar{y} = \frac{y_1 + y_2 + \dots + y_n}{n}$$

母集団Y

母比率  $p_y$



1または0の集まり

このとき  $\bar{x} - \bar{y}$  は期待値  $p_x - p_y$  分散  $\frac{p_x(1-p_x)}{m} + \frac{p_y(1-p_y)}{n}$  の正規分布になる

帰無仮説より  $p_x = p_y = p$  を代入して標準化



の分布は  $N(0, 1)$  (標準正規分布) (ただし  $m, n$  は大きな数)

母比率  $p_x = p_y = p$  だが  $p$  は未知なので

帰無仮説

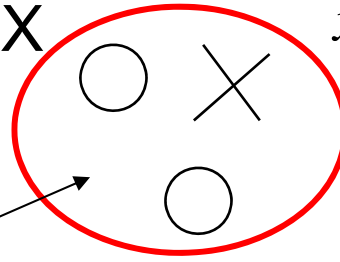
とする

有意水準  $\alpha$  を設定し、標本から得られた  $z$  が正規分布の棄却域に入るかどうか判定  
棄却域に入ったら帰無仮説を棄却する

# 比率の検定

(2) 異なる2つの母集団の間で、  
母比率に差があるといえるかどうか？

標本X



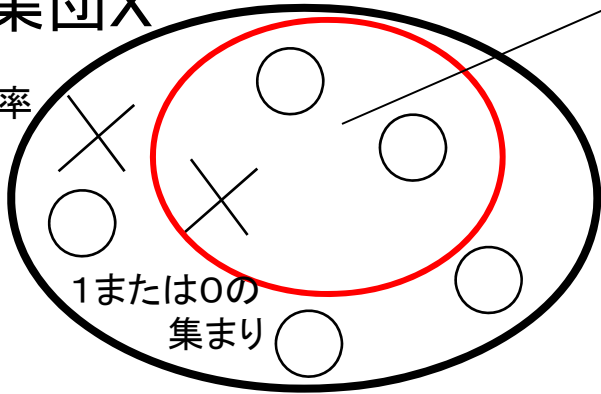
$x_1, x_2, \dots, x_m$

標本平均 (標本比率)

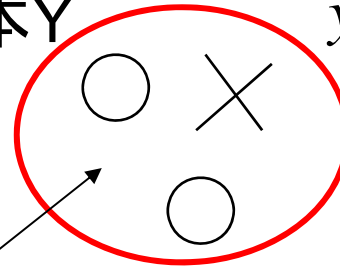
$$\bar{x} = \frac{x_1 + x_2 + \dots + x_m}{m}$$

母集団X

母比率  $p_x$



標本Y



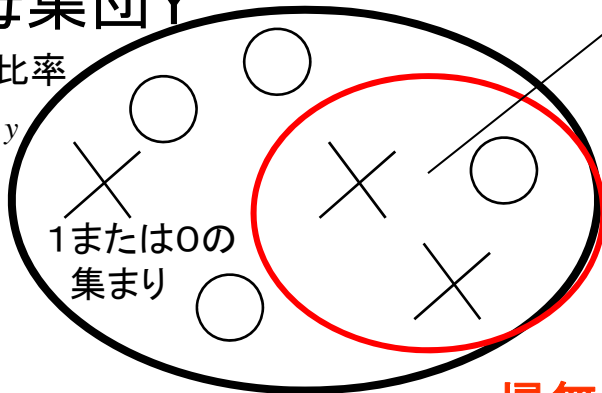
$y_1, y_2, \dots, y_n$

標本平均 (標本比率)

$$\bar{y} = \frac{y_1 + y_2 + \dots + y_n}{n}$$

母集団Y

母比率  $p_y$



このとき  $\bar{x} - \bar{y}$  は  
期待値  $p_x - p_y$  分散  
の正規分布になる

$$\frac{p_x(1-p_x)}{m} + \frac{p_y(1-p_y)}{n}$$

帰無仮説より  $p_x = p_y = p$  を代入して  
標準化

$$z = \frac{\bar{x} - \bar{y}}{\sqrt{p(1-p)\left(\frac{1}{m} + \frac{1}{n}\right)}}$$

の分布は  $N(0, 1)$   
(標準正規分布)

(ただし  $m, n$  は大きな数)

母比率  $p_x = p_y = p$

ただし  $p$  は未知なので

帰無仮説

$$p = \frac{m\bar{x} + n\bar{y}}{m+n} \text{ とする}$$

有意水準  $\alpha$  を設定し、標本から得られた  $z$  が正規分布の棄却域に入るかどうか判定  
棄却域に入ったら帰無仮説を棄却する

【練習問題】

43.3%

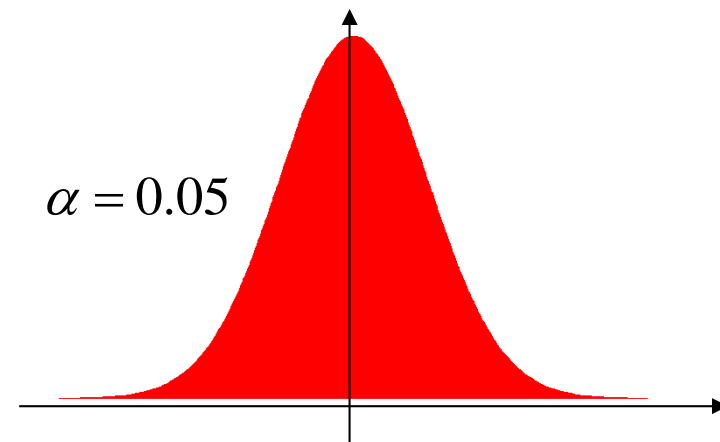
ある貨物船で鋼板を輸送したところ、60回の運行において積荷の鋼板における輸送ダメージのクレームが26件であった。そのため、ある装置を導入して対策を講じたところ、その後の40回の運行において積荷ダメージに対するクレームが14件であった。対策を講じた効果があったと判断すべきか？有意水準5%で検定せよ。

35%

ただし  $\frac{10}{12} = 0.833$  として計算せよ。

**帰無仮説:**  $p_x = p_y = p$   
(効果が無い)

**対立仮説:**  
(効果がある)



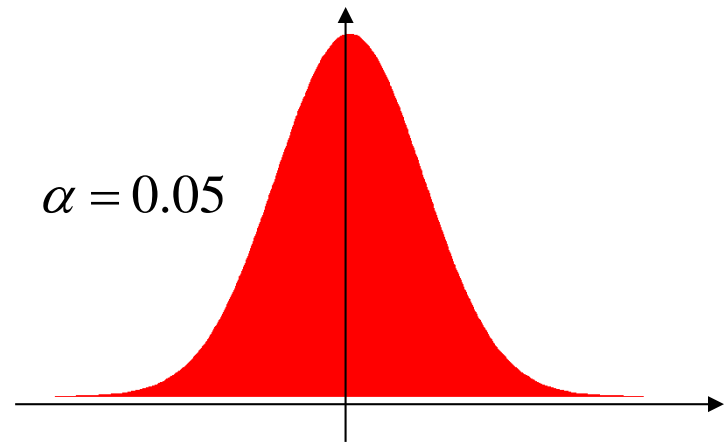
### 【練習問題】

ある貨物船で鋼板を輸送したところ、60回の運行において積荷の鋼板における輸送ダメージのクレームが26件であった。そのため、ある装置を導入して対策を講じたところ、その後の40回の運行において積荷ダメージに対するクレームが14件であった。対策を講じた効果があったと判断すべきか？有意水準5%で検定せよ。

ただし  $\frac{10}{12} = 0.833$  として計算せよ。

**帰無仮説:**  $p_x = p_y = p$      $p = \frac{26+14}{60+40} = 0.4$   
(効果が無い)

**対立仮説:**  $p_x > p_y$      $\alpha = 0.05$   
(効果がある)





## 【練習問題】

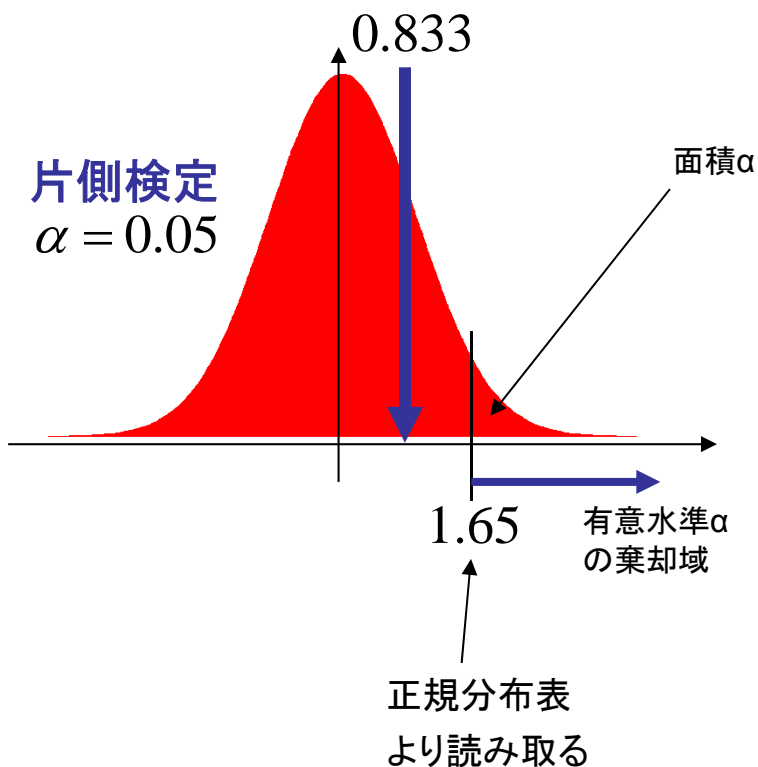
ある貨物船で鋼板を輸送したところ、60回の運行において積荷の鋼板における輸送ダメージのクレームが26件であった。そのため、ある装置を導入して対策を講じたところ、その後の40回の運行において積荷ダメージに対するクレームが14件であった。対策を講じた効果があったと判断すべきか？有意水準5%で検定せよ。

ただし  $\frac{10}{12} = 0.833$  として計算せよ。

**帰無仮説:**  $p_x = p_y = p$   $p = \frac{26+14}{60+40} = 0.4$   
(効果が無い)

**対立仮説:**  $p_x > p_y$   
(効果がある)

$$z = \frac{\bar{x} - \bar{y}}{\sqrt{p(1-p)\left(\frac{1}{m} + \frac{1}{n}\right)}}$$
$$= \frac{\frac{26}{60} - \frac{14}{40}}{\sqrt{0.4 \times (1-0.4) \left(\frac{1}{60} + \frac{1}{40}\right)}} = \frac{10}{12} \approx 0.833$$





## 【練習問題】

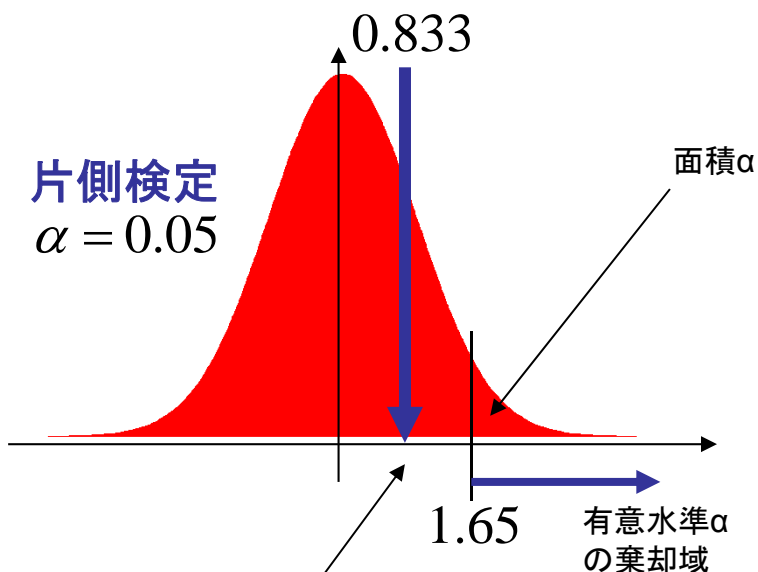
ある貨物船で鋼板を輸送したところ、60回の運行において積荷の鋼板における輸送ダメージのクレームが26件であった。そのため、ある装置を導入して対策を講じたところ、その後の40回の運行において積荷ダメージに対するクレームが14件であった。対策を講じた効果があったと判断すべきか？有意水準5%で検定せよ。

ただし  $\frac{10}{12} = 0.833$  として計算せよ。

**帰無仮説:**  $p_x = p_y = p$   $p = \frac{26+14}{60+40} = 0.4$   
(効果が無い)

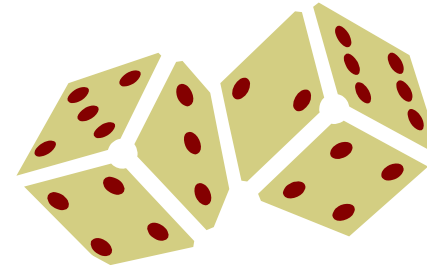
**対立仮説:**  $p_x > p_y$   
(効果がある)

$$z = \frac{\bar{x} - \bar{y}}{\sqrt{p(1-p)\left(\frac{1}{m} + \frac{1}{n}\right)}} = \frac{\frac{26}{60} - \frac{14}{40}}{\sqrt{0.4 \times (1-0.4) \left(\frac{1}{60} + \frac{1}{40}\right)}} = \frac{10}{12} \approx 0.833$$



観測されたデータは帰無仮説の分布の棄却域に入らないので、帰無仮説(効果が無い)は棄却できない  
=効果があるとはいえない

## 【レポート課題】 2018.05.15



レポート題目:「変形サイコロの検定」

- 1) 講義にて各自に配布した変形サイコロを振り、各目の出現回数について集計せよ。  
少なくとも100回以上サイコロを振ってデータを取得せよ。
- 2) 上記変形サイコロの、正方形端面(白い数字)の出る割合について、  
信頼係数 95% で区間推定せよ。
- 3) 上記変形サイコロが、立方体サイコロと違うかどうか検定したい。  
正方形端面(白い数字)の出る割合  $p = 1/3 = 0.3333$  を帰無仮説として  
有意水準 1% で検定せよ。



## 【注意事項】

- ・配布した変形サイコロは別の課題でも使用し、**使用後回収するので大切に保管すること。**
- ・10点満点で採点し期末試験の点に加える。試験に自信が無ければ必ず提出すること。
- ・レポートには**A4用紙片面**を使用し、**レポートの表題と学籍番号・氏名・サイコロ番号**を明記する。  
複数枚数で構成される場合は**ホチキスで留めること。**(散逸し採点できないため)  
A4用紙でないもの、複数枚がホチキス留めでないものは提出しても0点なので注意

**提出期限: 2018年 5月29日(火) 午後4時まで** 提出先: W2-6F 634号室

55.8%

【演習問題】

2018.05.15

学籍番号

氏名

---

ある造船所Aで作業員240名に聞き取り調査を行ったところ、作業中にヒヤリとした経験をした人が134名であった。別の造船所Bにおいて作業員160名に聞き取り調査を行ったところ、同様の回答をした人が106名であった。  
造船所AとBでは安全対策の効果に差があるといえるか？有意水準5%で検定せよ。

66.3%

【演習問題】

学籍番号

氏名

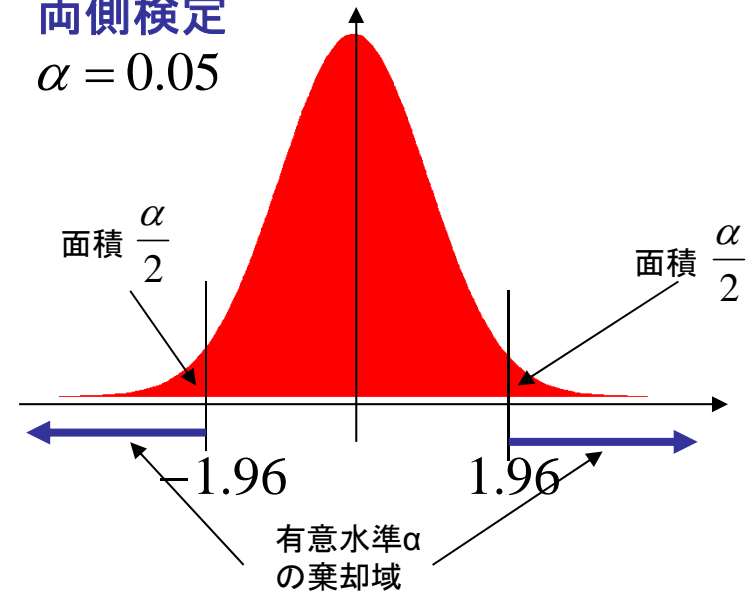
ある造船所Aで作業員240名に聞き取り調査を行ったところ、作業中にヒヤリとした経験をした人が134名であった。別の造船所Bにおいて作業員160名に聞き取り調査を行ったところ、同様の回答をした人が106名であった。  
造船所AとBでは安全対策の効果に差があるといえるか？有意水準5%で検定せよ。

帰無仮説:  $p_x = p_y = p$        $p = \frac{134+106}{240+160} = 0.6$

対立仮説:  $p_x \neq p_y$   $\longrightarrow$

$$z = \frac{\bar{x} - \bar{y}}{\sqrt{p(1-p)\left(\frac{1}{m} + \frac{1}{n}\right)}} = \frac{\frac{134}{240} - \frac{106}{160}}{\sqrt{0.6 \times (1-0.6)\left(\frac{1}{240} + \frac{1}{160}\right)}} = -\frac{50}{24} \approx -2.08$$

両側検定  
 $\alpha = 0.05$



【演習問題】

学籍番号

氏名

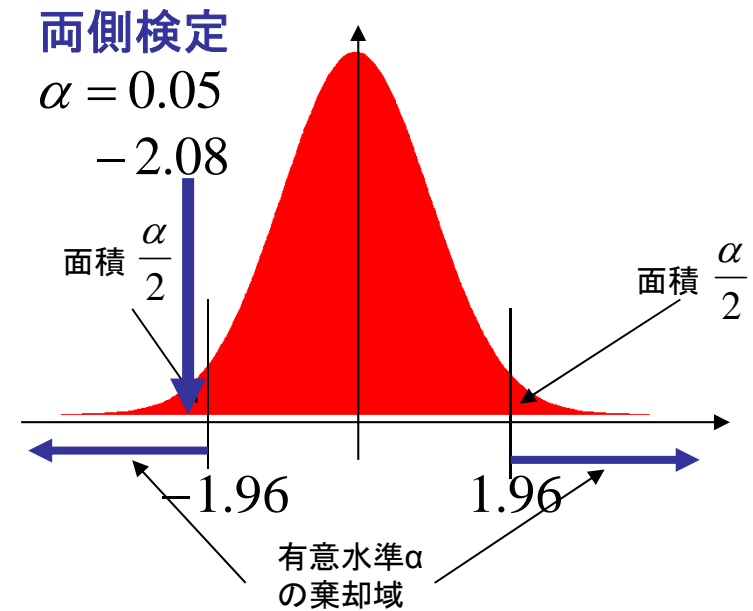
ある造船所Aで作業員240名に聞き取り調査を行ったところ、作業中にヒヤリとした経験をした人が134名であった。別の造船所Bにおいて作業員160名に聞き取り調査を行ったところ、同様の回答をした人が106名であった。  
造船所AとBでは安全対策の効果に差があるといえるか？有意水準5%で検定せよ。

帰無仮説:  $p_x = p_y = p$        $p = \frac{134+106}{240+160} = 0.6$

対立仮説:  $p_x \neq p_y$   $\longrightarrow$

$$z = \frac{\bar{x} - \bar{y}}{\sqrt{p(1-p)\left(\frac{1}{m} + \frac{1}{n}\right)}} = \frac{\frac{134}{240} - \frac{106}{160}}{\sqrt{0.6 \times (1-0.6)\left(\frac{1}{240} + \frac{1}{160}\right)}} = -\frac{50}{24} \approx -2.08$$

もし片側検定の場合、  
棄却域は  $1.65 < z$



棄却域に入るので、  
仮説は棄却される  
＝効果に差がある