6 基本統計量

基本統計量とは、データの分布の特徴を記述したり要約したりするために必要な指標です。データ表から、**(偏差)平方和 S** や**不偏分散 V**、**標準偏差 s** の計算ができるようにしておきましょう。

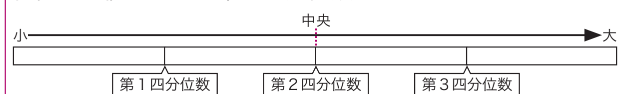
図表1.5 母集団とサンプルの関係

	母集団(母数)	サンプル(統計量)
平均	母平均 μ	平均値 $\bar{x}$
ばらつき		(偏差)平方和 S＊
	母分散 σ^2＊	不偏分散 V
	母標準偏差 σ	標準偏差 s

理解しておきたい基本統計量

基本統計量	概要
平均値 $\bar{x}$	個々のデータ（x_1, x_2, x_3, x_4, …, x_n）を全部足して合計を求め、その合計をデータの個数で割ったもの。 $\bar{x}=\frac{1}{n}(x_1+x_2+x_3+x_4,\cdots+x_n)=\frac{1}{n}\sum_{i=1}^{n}x_i$ ◉ $\sum_{i=1}^{n}x_i$ について、Σの上下にある n と $i=1$ は省略して表記されることがある（以下同） 〈例〉ある部品の長さ（単位：mm）について、 ５個のデータ　{3.0, 3.2, 3.3, 3.3, 3.7}の平均値 $\bar{x}$ は、 $\bar{x}=\frac{3.0+3.2+3.3+3.3+3.7}{5}=\frac{16.5}{5}=3.3$
中央値(メディアン) $\tilde{x}$	データを昇順(大きさの順)に並べた際に､中央に位置する値。 データの数が偶数の場合は､中央の２つの値の平均値をとる。 • 測定値が奇数個ある場合……中央に位置する値 〈例〉{４, ５, ６, ７, ８}の中央値は、 $\tilde{x}=6$ • 測定値が偶数個ある場合……中央の２つの値の平均値 〈例〉{５, ６, ７, ８}の中央値は、 $\tilde{x}=\frac{6+7}{2}=6.5$
最頻値(モード)	データの中で、最も多く現れる値。

＊ちなみに、平方和 S をデータの個数 $n-1$ で割った値を分散（s^2）という。基本統計量として「不偏分散」やその平方根である「不偏標準偏差」が多く用いられるのは、理論上の値より実測値（データ）の方が小さくなる傾向があり、それを補正するため。当然、データの個数nが大きくなればなるほど、理論上の値と実測値の差は小さくなっていく

範囲 R	データの最大値と最小値の差（R＝*Range*）。
偏差	個々のデータから平均値を引いた値。$x_n-\bar{x}$
(偏差)平方和 S	偏差(個々のデータから平均値を引いた値)を２乗し、これらを合計した値。 $S=\sum(x_i-\bar{x})^2=\sum x_i^2-2\sum x_i\cdot\bar{x}+\sum\bar{x}^2$ $=\sum x_i^2-2\bar{x}\sum x_i+\bar{x}^2\sum 1$ $=\sum x_i^2-2\frac{\sum x_i}{n}\cdot\sum x_i+n\left(\frac{\sum x_i}{n}\right)^2$ $=\sum x_i^2-\frac{(\sum x_i)^2}{n}$ 〈例〉{２, ３, ４, ５, ６}の平方和 S は、 $S=(4+9+16+25+36)-\frac{(2+3+4+5+6)^2}{5}$ $=90-\frac{20\times 20}{5}=90-80=10$
不偏分散 V※	(偏差)平方和を自由度 ϕ（サンプルサイズ $n-1$）で割った値。母分散の推定値として用いられる。 $V=\frac{S}{\phi}=\frac{S}{n-1}=\frac{\sum(x_i-\bar{x})^2}{n-1}=\frac{\sum x_i^2-\frac{(\sum x_i)^2}{n}}{n-1}$ 〈例〉平方和 $S=15$、観測個数６個のときの不偏分散 V は、 $V=\frac{15}{6-1}=\frac{15}{5}=3$
(不偏)標準偏差 s	不偏分散の平方根（分散の値の次元をデータと同じ次元に落とした値）。 $s=\sqrt{V}=\sqrt{\frac{S}{n-1}}=\sqrt{\frac{\sum(x_i-\bar{x})^2}{n-1}}=\sqrt{\frac{\sum x_i^2-\frac{(\sum x_i)^2}{n}}{n-1}}$ 〈例〉不偏分散 $V=9$ のときの標準偏差 s は、 $s=\sqrt{9}=3$
変動係数 CV	標準偏差を平均値で割った値。$CV=s/\bar{x}$ 〈例〉標準偏差 $s=0.6$、平均値 $\bar{x}=3.0$ のときの変動係数 CV は、 $CV=\frac{0.6}{3.0}=0.2$
四分位数	データを大きさの順に並べたとき、４等分する位置の値。４等分すると、仕切りが３つできることになる。１番目の仕切りに位置する値のことを第１四分位数、２番目の仕切りに位置する値のことを第２四分位数(中央値)、３番目の仕切りに位置する値のことを第３四分位数という。  データの数が偶数の場合は、中央値と同様に、中央の２つの値の平均値をとる。

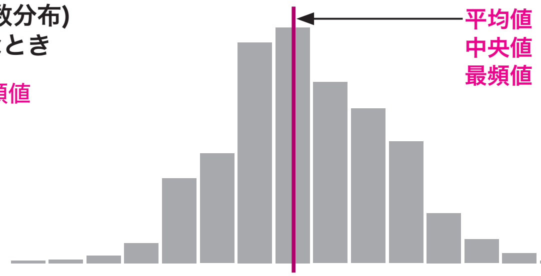
※(不偏)分散と平均平方は同義であるが、過去問の出題実績に則り、第Ｉ章(データの取り方とまとめ方)では｢(不偏)分散｣と表記し、第５章(単回帰分析)においては｢平均平方｣と表記する

基本統計量において、**平均値**、**中央値(メディアン)**、**最頻値(モード)**の関係にはいくつかのパターンがあります。**出題された**こともありますので、理解しておきましょう。

図表1.6　左右対称なヒストグラム

①ヒストグラム(度数分布)がほぼ左右対称なとき

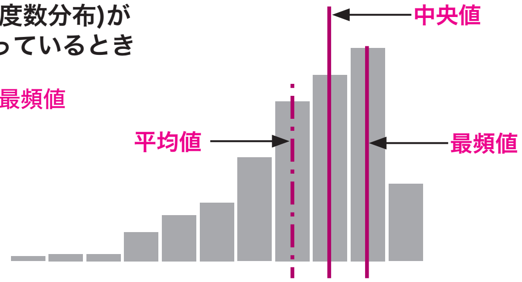
平均値≒中央値≒最頻値



図表1.7　右に偏ったヒストグラム

②ヒストグラム(度数分布)が右のほうに偏っているとき

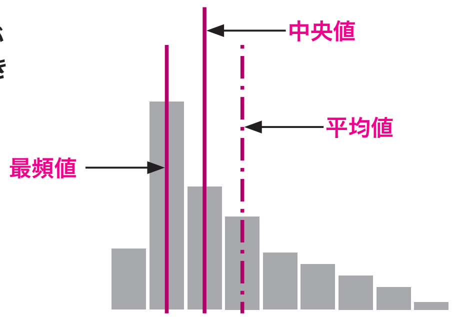
平均値＜中央値＜最頻値



図表1.8　左に偏ったヒストグラム

③ヒストグラム(度数分布)が左のほうに偏っているとき

平均値＞中央値＞最頻値



練習問題　　赤シートで正解を隠して設問に答えてください(解説はP.20)。

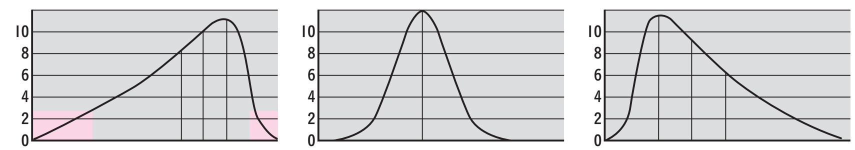
【問１】　下記のデータにおいて、平均値、中央値、最頻値、範囲、平方和、分散、標準偏差、変動係数および第３四分位数を求めよ。

{2, 3, 4, 4, 4, 5, 6, 7, 7, 8}

正解　平均値5　中央値4.5　最頻値4　範囲6　平方和34
分散3.78　標準偏差1.94　変動係数0.39　第３四分位数7

【問２】　下記の３つのデータ分布グラフにおいて、平均値、中央値、最頻値の関係を正しく表した式を選べ。

①　　②　　③



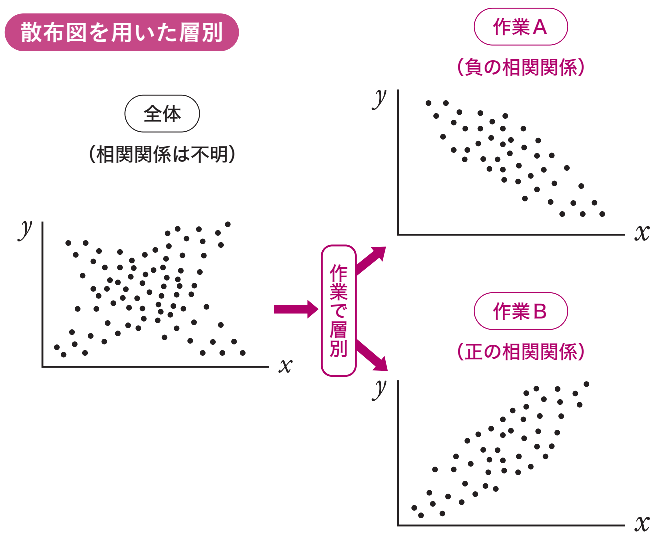
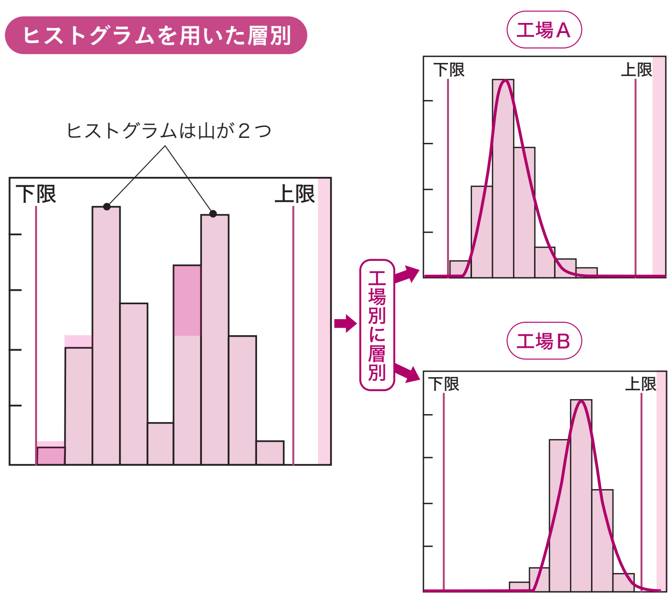
ア．平均値＝中央値＝最頻値
イ．最頻値＜中央値＜平均値
ウ．平均値＜中央値＜最頻値

正解　①ウ　②ア　③イ

【問３】　下記の(１)～(３)に入る語句を答えよ。

- 母集団のサンプリング単位が何らかの順序で並んでいる際に、一定間隔でサンプリング単位を取るサンプリングを（１）という。
- 母集団をいくつかの部分母集団に分割して、各部分母集団から１つ以上のサンプリング単位を取るサンプリングを（２）という。
- 必要な数のサンプリング単位が母集団から一度に取られるサンプリング又は母集団に戻すことなく繰り返し取られるサンプリングを（３）という。

正解　（１）系統サンプリング　（２）層別サンプリング
（３）非復元サンプリング

層別	 

練習問題　　赤シートで正解を隠して設問に答えてください（解説はP.33から）。

【問１】　以下のＡ群にある新ＱＣ七つ道具において、それぞれ該当する説明（Ｂ群）を挙げよ。

〈Ａ群〉
- **ア**．系統図法
- **イ**．アローダイアグラム法
- **ウ**．親和図法

〈Ｂ群〉
- **ア**．言語データを、イメージの類似性の観点でまとめていく。
- **イ**．進行上の順序関係を明確にして、その順序に沿った工程計画を作成する。
- **ウ**．ゴールを明確に設定し、そこに至る手段・方策を系統づけて展開する。

正解　（Ａ群）ア⇔（Ｂ群）ウ　（Ａ群）イ⇔（Ｂ群）イ　（Ａ群）ウ⇔（Ｂ群）ア

【問２】　下記の（１）～（５）に当てはまる語句を選択肢から選べ。

①連関図とは、原因と［（１）］、目的と［（２）］などが複雑に絡み合った問題の関係を、論理的につなぐことによって問題解明を図るものである。
②ＰＤＰＣ法とは、［（３）］を実施するうえで、トラブル防止やさまざまな結果の予測を行うことで、［（４）］をできるだけ望ましい方向に導くものである。
③［（５）］法とは、行と列に属する要素により構成された二元表の交点に着目して、その二元的関係の中から問題解決への着想を得るものである。

〈選択肢〉
ア．計画　**イ**．目標　**ウ**．課題　**エ**．分析　**オ**．手段　**カ**．結果
キ．プロセス　**ク**．マトリックス・データ解析
ケ．マトリックス図　**コ**．パレート図

正解　（１）カ　（２）オ　（３）ア　（４）キ　（５）ケ

〈例〉サイコロを６回振って１の目がx回出るときの確率は二項分布に従う。そこで$x=4$の場合、$P(4)$の確率は次のように求めることができる。

１の目が出る確率は、$\frac{1}{6}$。１の目が出ない確率は、$\frac{6}{6}-\frac{1}{6}=\frac{5}{6}$。

$$P(4)={}_6C_4\times\left(\frac{1}{6}\right)^4\times\left(\frac{5}{6}\right)^2$$

$$=\frac{6\times5\times4\times3\times2\times1}{4\times3\times2\times1\times2\times1}\times\frac{1^4}{6^4}\times\frac{5^2}{6^2}$$

$$=\frac{6\times5}{2\times1}\times\frac{1}{1296}\times\frac{25}{36}\ \fallingdotseq 0.008=0.8\%$$

4 ポアソン分布

ポアソン分布とは、**ある時間間隔で発生する事象**の回数を表す離散確率分布です。

非負整数値（０，１，２，…）をとる離散的な確率変数をxとし、単位時間当たりの事象の平均発生回数をλ（ラムダ）とし、事象が発生する回数をkとすると、

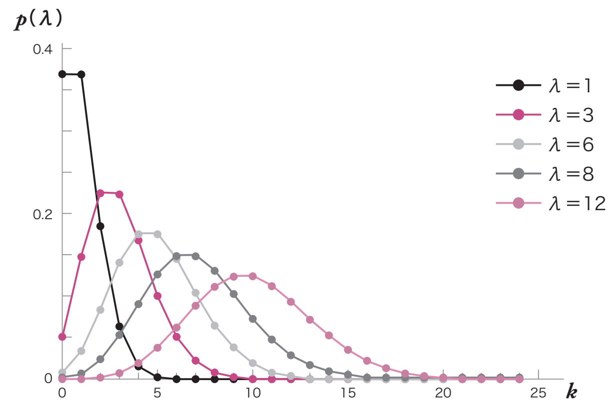
$$P(x=k)=\frac{\lambda^k e^{-\lambda}}{k!}$$　※e：自然対数の底（2.718…）

と表され、確率変数xは母数λの**ポアソン分布**に従うといいます。

ポアソン分布は、二項分布$B(n,\ p)$において、npを一定の値λとし、このλを一定に保った状態で、nを十分に大きくして、pを十分に小さくした場合（調査数nが相当に多い中、不良品率pは極めて小さい場合）の確率分布であり、確率の極めて小さい事象が、多数回の試行の結果生じるものととらえることができ、**まれな現象の確率分布**ともいわれます。

なお、**$\lambda\geqq5$であれば、ポアソン分布を実用上、正規分布として扱ってよい**とされています。図表３．３を見ると、$\lambda=6,8,12$のグラフが正規分布（左右対称の形）になっていることがわかります。

図表３．３ ポアソン分布のグラフの一例



〈例〉ある板１枚当たりのへこみの数が一定の単位中に現れる欠点数の確率がポアソン分布に従うとき、へこみの平均箇所数が４か所である場合に、へこみが１か所もない確率と、へこみが１か所ある確率を求める（ただし、$e^{-4}=0.0183$とする）。

- へこみが１か所もない確率は、$\lambda=4$、$k=0$の場合。

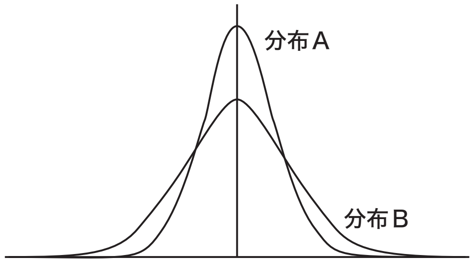
$$P(0)=\frac{4^0e^{-4}}{0!}=e^{-4}=0.0183$$

- へこみが１か所ある確率は、$\lambda=4$、$k=1$の場合。

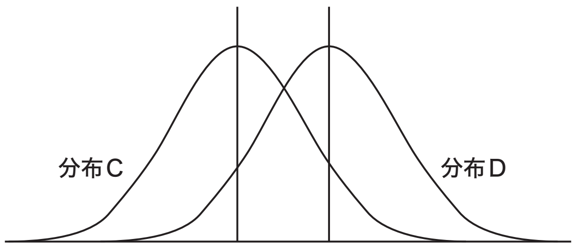
$$P(1)=\frac{4^1e^{-4}}{1!}=4\times e^{-4}=0.0732$$

5 期待値、分散、共分散の基本性質（公式）

期待値（Expected value）とは、確率変数が取る値を、確率によって重み付けした**平均値**のことです。一方、分散（Variance）とは、確率変数のばらつき具合を表し、確率変数と期待値の差を２乗したものに、確率で重みをつけた重み付き**算術平均**となります。



上図において、分布Aと分布Bを比較すると、分布Aと分布Bの期待値(＝平均値)は同じですが、分布Aの分散(＝ばらつき)は分布Bの分散に比べて小さいといえます。



上図において、分布Cと分布Dを比較すると、分布Cと分布Dの分散(＝ばらつき)は同じですが、分布Cの期待値(＝平均値)は分布Dの期待値に比べて小さいといえます。

なお、**共分散**(Covariance)とは、２つの確率変数の関連性を測る尺度として用いられ、大きさが同じ２つのデータの間での、平均からの偏差の積の平均値を表します。

- $Cov(X,\ Y) = E\{(X - E(X))(Y - E(Y))\}$
  $\qquad\qquad = E(XY) - E(X)E(Y)$

	離散型確率変数	連続型確率変数
期待値 E	$E(X) = \sum_i X_i f(X_i)$	$E(X) = \int_{-\infty}^{\infty} Xf(X)dx$
分　散 V	$V(X) = E\{(X - E(X))^2\}$	

※共分散の期待値 E と分散 V の計算式は上記の通り

期待値 E、分散 V、共分散 Cov には、それぞれ以下の基本性質(公式)があります(X, Y は確率変数、a、b は定数を表す)。

〈期待値 E〉

- $E(a) = a$
- $E(aX) = aE(X)$
- $E(X + a) = E(X) + a$
- $E(X + Y) = E(X) + E(Y)$
- $E(X - Y) = E(X) - E(Y)$
- $E(aX + Y + b) = aE(X) + E(Y) + b$

※ X と Y が無相関(互いに独立)の場合、

- $E(XY) = E(X)E(Y)$

期待値と分散の基本性質

❶確率変数 x に定数 a を加えると、期待値 E は a だけ増すが、分散 V は変わらない。
❷確率変数 x に定数 a をかけると、期待値 E は元の a 倍になるが、分散 V は a^2 倍になる。
❸２つの確率変数 x と y の和(差)の期待値 E は各確率変数の期待値の和(差)に等しい。
❹２つの独立した確率変数 x と y の和の分散 V は、各確率変数の分散の和に等しい。

〈分散 V〉

- $V(a) = 0$
- $V(aX) = a^2V(X)$
- $V(X + a) = V(X)$
- $V(X + aY) = V(X) + a^2V(Y)$
- $V(aX + Y + b) = a^2V(X) + V(Y)$
- $V(X + Y) = V(X) + V(Y) + 2Cov(X,\ Y)$
- $V(X - Y) = V(X) + V(Y) - 2Cov(X,\ Y)$

$\Big\}$ X と Y が互いに独立ではない場合

〈共分散 Cov〉

- $Cov(X,\ Y) = E\{(X - E(X))(Y - E(Y))\} = E(XY) - E(X)E(Y)$

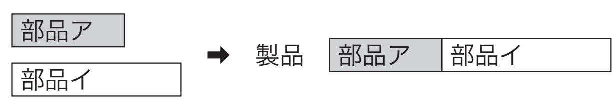
※ X と Y が無相関(互いに独立)の場合、X と Y を合わせたものの分散は、X の分散と Y の分散を足した値になる

- $V(X + Y) = V(X) + V(Y)$
- $V(X - Y) = V(X) + V(Y)$

➡分散の加法性と呼ぶ。

〈例〉部品アの長さは母平均６㎝、母標準偏差0.4㎝であり、部品イの長さは母平均９㎝、母標準偏差0.6㎝である。部品アと部品イを横につなげて製品をつくるとき、その製品の長さの母平均と母標準偏差を求める。

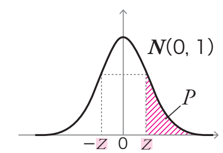
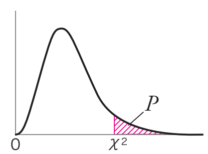
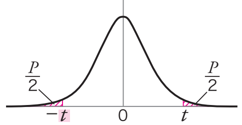
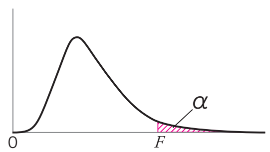
- 母平均＝E(ア)＋E(イ)＝６＋９＝15㎝
- 母標準偏差＝$\sqrt{\text{母分散}} = \sqrt{\sigma(\text{ア})^2 + \sigma(\text{イ})^2} = \sqrt{0.4^2 + 0.6^2} = \sqrt{0.52} \fallingdotseq$ **0.72**㎝



6 統計量の分布

正規分布をする母集団$N(\mu, \ \sigma^2)$について、仮説検定を行う際に用いる主な統計量の分布には、**標準正規分布**、**t分布**、**χ^2分布**、**F分布**があり、下の図表のとおり、使い分けます。

図表3．4　仮説検定に用いる主な統計量の分布

	平均に関する検定	ばらつき（分散）に関する検定
母分散が既知の場合	標準正規分布（Z分布）  	χ^2分布 
母分散が未知の場合	t 分布  	F 分布 

※図中のP、αは棄却域（有意水準）を示す

1）標準正規分布（Z分布）

母平均μ、母分散σ^2の母集団から大きさnのサンプル（標本）をランダムに抽出したとき、n個のサンプルの平均値$\bar{x}$の平均値（期待値）$E(x)$と分散$V(x)$は、

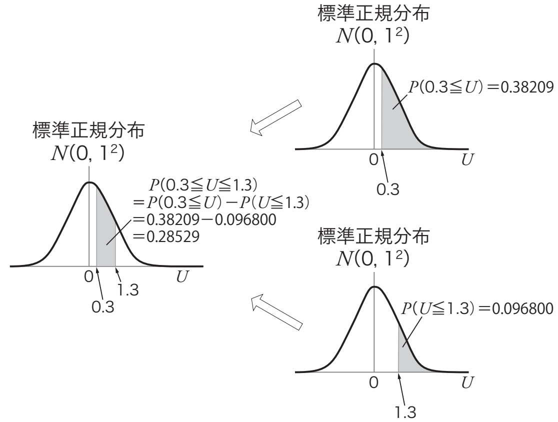
- $E(x)=\mu$
- $V(x)=\dfrac{\sigma^2}{n}$

となります。

また、$Z=(\bar{x}-\mu)/\sqrt{(\sigma^2/n)}$と規準化すると、**$Z$は$N(0, 1^2)$の標準正規分布に従います**。この$Z$を**検定統計量**といいます。

※実際の試験でも、分数の表記は例えばσ^2/nと$\frac{\sigma^2}{n}$が混在しているので要注意。

このnが十分に大きいと、以下の①大数の法則および②中心極限定理が成立します。例えば、標準正規分布$N(0, 1^2)$において、0.3以上1.3以下となる確率を求めると、0.28529となります（下の図表参照）。

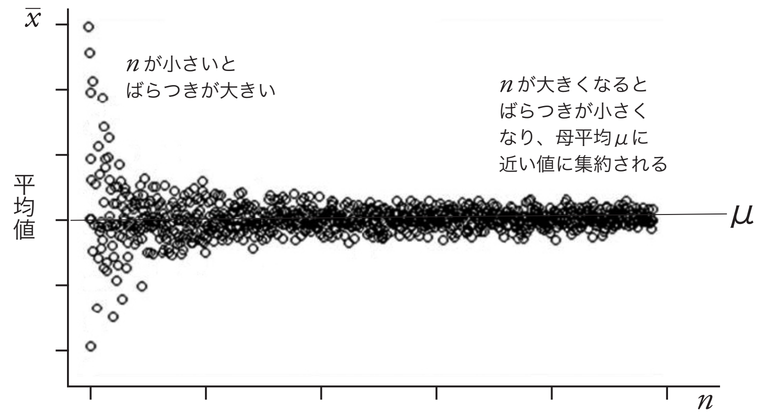


（Ⅰ）K_PからPを求める表

K_P	0.00	0.01	0.02
0.0	.50000	.49601	.49202
0.1	.46017	.45620	.45224
0.2	.42074	.41683	.41294
0.3	.38209	.37828	.37448
0.4	.34458	.34090	.33724
0.5	.30854	.30503	.30153
0.6	.27425	.27093	.26763
0.7	.24296	.23885	.23576
0.8	.21186	.20997	.20611
0.9	.18406	.18141	.17879
1.0	.15866	.15625	.15386
1.1	.13567	.13350	.13136
1.2	.11507	.11314	.11123
1.3	.096800	.095098	.093418
1.4	.080757	.079270	.077804
1.5	.066807	.065522	.064255

①大数の法則

標本平均$\bar{x}$は、**母平均μ**に近い値をとります。



②中心極限定理

母集団の分布がどんな分布であっても、標本の大きさを大きくしたときは**近似的に正規分布に従う**というものです。母集団の従う確率分布に関係なく、標本平均は期待値μ、分散σ^2/nの正規分布$N(\mu, \ \sigma^2/n)$に従うとみなせます。

2) t分布

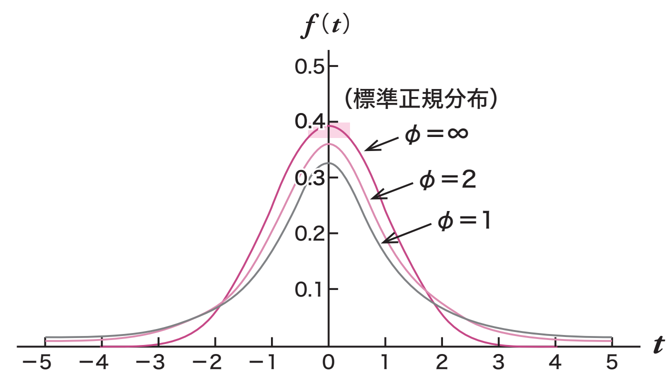
t分布は標準正規分布とよく似た形の分布で、パラメータである**自由度ϕ**によって分布の形が変わるという特徴を持っており、自由度ϕが大きくなるにつれて、標準正規分布に近づきます(下図参照)。

母分散が未知の値であり、n＝20 個程度のデータしか集められない場合に、母平均の値を推定したいときに利用される確率分布です。標準正規分布における検定統計量の式$Z=(\bar{x}-\mu)/\sqrt{(\sigma^2/n)}$において、$\sigma^2$の代わりに点推定量$\hat{\sigma}^2=V$を代入します($\hat{\sigma}^2$は$\sigma^2$の推定量)。

$N(\mu,\ \hat{\sigma}^2/n)$から$n$個のサンプルをとり、次式で与えられる**検定統計量tは、自由度$\phi=n-1$**のt分布となります。

検定統計量$t_0=\dfrac{\bar{x}-\mu}{\sqrt{V/n}}$、自由度$\phi$の$t$分布の**両側確率**$\alpha$の点を$t(\phi,\ P)$で表します。巻末の付表2($t$表)を用いて、例えば$\phi=5$, $P=0.05$の場合、$t(5, 0.05)=2.571$を読み取ることができます。

図表3.5 自由度ϕのt分布



〈例〉母平均μ＝3.0の正規母集団から大きさn＝7個のサンプルをとり、次の値を得た(母分散は未知とする)。サンプルの平均値は$\bar{x}$＝3.3、不偏分散はV＝0.3である。このときの検定統計量t_0は、

$$t_0=\frac{3.3-3.0}{\sqrt{\frac{0.3}{7}}}\fallingdotseq 1.45 \quad \text{となる。}$$

3) χ^2分布(カイの2乗分布)

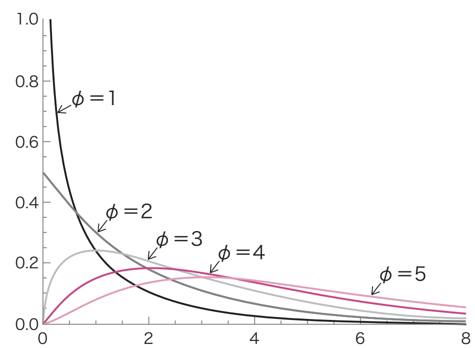
χ^2(カイ)分布は、**ばらつき(母分散σ^2)に関する検定と推定**に用います。

χ^2は、平方和Sを母分散σ^2で割ったものであり、$N(\mu,\ \sigma^2)$からn個のサンプルをとり、その平方和Sをσ^2で割ったものは自由度$\phi=n-1$のχ^2分布となります。

$$\chi^2=\frac{S}{\sigma^2}$$

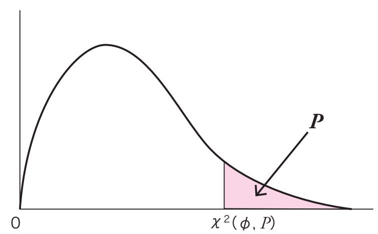
χ^2分布は**自由度ϕ**によって定まります。自由度ϕのχ^2の**上側確率**αの点を$\chi^2(\phi,\ P)$で表します。巻末の付表3(χ^2表)を用いて、例えば、$\phi=5$, $P=0.05$の場合、$\chi^2(5, 0.05)=11.0705$を読み取ることができます。

図表3.6 χ^2分布の確率密度関数



※**上側確率・下側確率**
ある確率分布において、確率変数が「ある値」以上になる確率を上側確率といい、「ある値」以下になる確率を下側確率という

図表3.7 自由度ϕのχ^2分布



〈例〉母標準偏差 σ =0.5の正規母集団から大きさ n = 6 個のサンプルをとり、次の値を得た。サンプルの平均値 $\bar{x}$ =3.5、平方和 S =2.5である。このときの統計値 χ_0^2 は、

$$\chi_0^2=\frac{2.5}{0.5^2}=10.0$$ となる。

4) F 分布

分散が等しい 2 つの正規分布 $N(\mu_1,\ \sigma^2)$、$N(\mu_2,\ \sigma^2)$ からそれぞれランダムに取られた n_1、n_2 のサンプルから得られた不偏分散をそれぞれ V_1、V_2 とすると、$F = V_1 / V_2$ は、自由度 $\phi_1 = n_1 - 1$、$\phi_2 = n_2 - 1$ の F 分布に従います。

自由度 ϕ_1、ϕ_2 の F 分布の**上側確率** α の点を、$F(\phi_1,\ \phi_2;\alpha)$ で表すと、F 分布の**下側確率** $\alpha-1$ の点は、$(\phi_1,\ \phi_2;\alpha)$ と表され、次式から求められます。

$$F(\phi_1,\ \phi_2;1-\alpha)=\frac{1}{F(\phi_2,\ \phi_1;\alpha)}$$

F 分布は、**分散分析や、正規分布に従う２つの母集団について「標準偏差が等しい」という仮説の検定など**に用います。

〈例〉ある正規母集団から大きさ n = 8 個のサンプルをとったところ、不偏分散 V_1 =0.50であった。さらに10個のサンプルをとったところ、V_2 =0.40であった。このときの F 検定統計値 F_0 は、

$$F_0=\frac{V_1}{V_2}=\frac{0.50}{0.40}=1.25$$ となる。

なお、F_0 を求めるときは、F_0 は 1 より大きくなるように、V_1 と V_2 のうち大きい方を分子とする必要がある。

練習問題

赤シートで正解を隠して設問に答えてください(解説はP.52から)。

【問１】 以下のＡ群にある３つの確率分布において、それぞれ該当する説明(Ｂ群)を挙げよ。

〈Ａ群〉

ア　正規分布

イ　二項分布

ウ　ポアソン分布

〈Ｂ群〉

ア　計数値データにおいて、ベルヌーイ試行を複数回行って、成功する回数が従う確率分布を表す。

イ　平均値と分散によって定まる分布で、調査対象数を増やしていくと、左右対称の釣鐘型の分布となる。

ウ　計数値データにおいて、まれに起こる現象の出現度数分布である。

正解　(Ａ群)ア⇔(Ｂ群)イ　(Ａ群)イ⇔(Ｂ群)ア　(Ａ群)ウ⇔(Ｂ群)ウ

【問２】 下記の(ａ)、(ｂ)に入る語句を答えよ。

$N(50,\ 2^2)$ の正規分布に従っている、薬品の入ったボトルの容量(単位：mL)を生産している工場において、100,000本のボトルを生産した。

容量が48mL未満となるボトルの数は［（ａ）］となる。

また、48mL以上かつ53mL未満の範囲にあるボトルを適合品とする場合、不適合品となるボトルの数は［（ｂ）］となる。

正解　（ａ）15,866 本　（ｂ）22,547 本

【問３】 下記の(ａ)、(ｂ)に入る語句を答えよ。

母不適合品率５％の工程で、10個のサンプルを採取したとき、P =0.05, n =10のときの二項分布(次ページのグラフ)に従うとすると、不適合品が含まれない確率は［（ａ）］%で、不適合品が３個以下となる確率は［（ｂ）］%である。

図表4.3　検定と推定の種類

扱うデータ	検定と推定の種類	統計量計算に用いる指標	検定統計量の計算式	検定に用いる分布
計量値	1つの母平均（母分散が未知の場合）	データの平均値 $\bar{x}$ 分散 V	$t_0=\dfrac{\bar{x}-\mu_0}{\sqrt{\dfrac{V}{n}}}$	t 分布
	1つの母平均（母分散が既知の場合）	データの平均値 $\bar{x}$ 母分散 σ^2	$Z_0=\dfrac{\bar{x}-\mu_0}{\sqrt{\dfrac{\sigma^2}{n}}}$	標準正規分布
	1つの母分散	母分散 σ^2 平方和 S	$\chi^2{}_0=\dfrac{S}{\sigma_0{}^2}$	χ^2分布
	2つの母平均の差（母分散が既知の場合）	データの平均値 $\bar{x}$ 母分散 σ^2	$Z_0=\dfrac{\bar{x}_1-\bar{x}_2-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma_1{}^2}{n_1}+\dfrac{\sigma_2{}^2}{n_2}}}$	標準正規分布
	2つの母平均の差（母分散が未知で、母分散が等しいと考えられる場合）	データの平均値 $\bar{x}$ 分散 V	$t_0=\dfrac{\bar{x}_1-\bar{x}_2}{\sqrt{V\left(\dfrac{1}{n_1}+\dfrac{1}{n_2}\right)}}$	t 分布
	2つの母平均の差（母分散が未知で、母分散が等しいかどうかわからない場合）	データの平均値 $\bar{x}$ 分散 V	$t_0=\dfrac{\bar{x}_1-\bar{x}_2}{\sqrt{\dfrac{V_1}{n_1}+\dfrac{V_2}{n_2}}}$	t 分布 （ウェルチの(t)検定）
	2つの母分散の比	分散 V	$F_0=\dfrac{V_1}{V_2}$（$V_1\geqq V_2$の場合） $F_0=\dfrac{V_2}{V_1}$（$V_1<V_2$の場合）	F 分布
	データに対応がある場合	母平均の差の平均値 $\bar{d}$ 分散 V_d	$t_0=\dfrac{\bar{d}}{\sqrt{\dfrac{V_d}{n}}}$	t 分布 母平均の差の平均値と分散による計算
計数値	母不適合品率	不適合品率 $P=\dfrac{x}{n}$	$u_0=\dfrac{x-nP_0}{\sqrt{nP_0(1-P_0)}}$	二項分布
	2つの母不適合品率の違い	不適合品率 $P=\dfrac{x}{n}$	$u_0=\dfrac{P_1-P_2}{\sqrt{\bar{P}(1-\bar{P})\left(\dfrac{1}{n_1}+\dfrac{1}{n_2}\right)}}$	
	母不適合数（母欠点数）	不適合数 c （欠点数）	$u_0=\dfrac{\bar{c}-m_0}{\sqrt{\dfrac{m_0}{n}}}$	ポアソン分布 （$m\geqq 5$で正規分布に近似できる）
	2つの母不適合数(母欠点数)の違い	不適合数 c （欠点数）	$u_0=\dfrac{c_1-c_2}{\sqrt{\bar{c}\left(\dfrac{1}{n_1}+\dfrac{1}{n_2}\right)}}$	
	複数の母集団での不適合品率の比較等	不適合品率	期待度数 $t_{ij}=T_{i.}\times T_{.j}/T_{..}$ $T_{i.}$：分割表の列の合計値、 $T_{.j}$：分割表の行の合計値、 $T_{..}$：分割表の合計値	分割表

①σ^2が未知の場合

手順1 仮説の設定

- 帰無仮説H_0：$\mu=\mu_0$
- 対立仮説H_1：$\mu\neq\mu_0$（両側検定）、
  $\mu>\mu_0$（右片側検定）、$\mu<\mu_0$（左片側検定）のいずれか※
  ※与条件などによる

手順2 有意水準αの設定

通常は$\alpha=0.05$とする。

手順3 統計量の計算

データの平均値$\bar{x}$および分散Vにより、t_0を求める。

$$t_0=\frac{\bar{x}-\mu_0}{\sqrt{\frac{V}{n}}}$$

手順4 検定

手順3で求めた検定統計量と確率分布表から求めた棄却限界値とを比較して判定する。

- H_1：$\mu\neq\mu_0$の場合、$|t_0|$※$\geqq t(\phi,\ \alpha)$であれば有意となり、H_0：$\mu=\mu_0$は棄却される。すなわちH_1が成立する。
- H_1：$\mu>\mu_0$の場合、$t_0\geqq t(\phi,\ 2\alpha)$であれば有意となり、$H_0$：$\mu=\mu_0$は棄却される。すなわち$H_1$が成立する。
- H_1：$\mu<\mu_0$の場合、$t_0\leqq -t(\phi,\ 2\alpha)$であれば有意となり、$H_0$：$\mu=\mu_0$は棄却される。すなわち$H_1$が成立する。

※$|t_0|$はt_0の絶対値

手順5 推定（点推定、区間推定）

- 点推定　：$\hat{\mu}=\bar{x}$
- 区間推定：$\bar{x}-t(\phi,\ \alpha)\sqrt{\frac{V}{n}}\leqq\mu\leqq\bar{x}+t(\phi,\ \alpha)\sqrt{\frac{V}{n}}$

〈例〉平均値が大きくなったか否かの検定（母集団の分散が未知の場合）

母平均：$\mu_0=9.0$の母集団からデータ数：$n=10$のサンプルをとった結果、標本平均値：$\bar{x}=10.0$、不偏分散：$V=0.4$であった。このとき、平均値が大きくなったかどうかの検定の手順は以下のとおりである。

手順1 仮説の設定

- 帰無仮説H_0：$\mu=\mu_0$
- 対立仮説H_1：$\mu>\mu_0$（右片側検定）

手順2 有意水準αの設定

$\alpha=0.05$とする

手順3 統計量の計算

$$t_0=\frac{\bar{x}-\mu_0}{\sqrt{V/n}}=\frac{10.0-9.0}{\sqrt{\frac{0.4}{10}}}=\frac{1.0}{\sqrt{0.04}}=5.0$$

手順4 検定

$\mu>\mu_0$　→　右片側検定

t表より、$\alpha=0.05$、$\phi=n-1=10-1=9$の棄却限界値tを求めると、$t(\phi,\ 2\alpha)=t(9,\ 0.10)=1.833$となる。※$t$表は両側検定の表となっているため、片側検定の場合はαの値を2倍にする。

$t_0=5.0>t=1.833$となることから、有意となる。帰無仮説は棄却され、対立仮説が成立する。

（参考）t表（一部抜粋）

ϕ ＼ P	0.50	0.40	0.30	0.20	0.10	0.05
1	1.000	1.376	1.963	3.078	6.314	12.706
2	0.816	1.061	1.386	1.886	2.920	4.303
3	0.765	0.978	1.250	1.638	2.353	3.182
4	0.741	0.941	1.190	1.533	2.132	2.776
5	0.727	0.920	1.156	1.476	2.015	2.571
6	0.718	0.906	1.134	1.440	1.943	2.447
7	0.711	0.896	1.119	1.415	1.895	2.365
8	0.706	0.889	1.108	1.397	1.860	2.306
9	0.703	0.883	1.100	1.383	1.833	2.262
10	0.700	0.879	1.093	1.372	1.812	2.228

手順5 **推定**

● 点推定：$\hat{\mu}=\bar{x}=10.0$

● 区間推定：$\bar{x}-t(\phi,\ \alpha)\sqrt{\dfrac{V}{n}} \leqq \mu \leqq \bar{x}+t(\phi,\ \alpha)\sqrt{\dfrac{V}{n}}$

（上限）　$\mu_U=\bar{x}+t(9,0.05)\sqrt{\dfrac{V}{n}}=10.0+2.262\sqrt{\dfrac{0.4}{10}}$

$=10.0+0.4524=10.4524$

（下限）　$\mu_L=\bar{x}-t(9,0.05)\sqrt{\dfrac{V}{n}}=10.0-2.262\sqrt{\dfrac{0.4}{10}}$

$=10.0-0.4524=9.5476$

よって、$9.5476 \leqq \mu \leqq 10.4524$

②σ^2が既知の場合（手順２までは①と同じ）

手順1 **仮説の設定**

● 帰無仮説H_0：$\mu=\mu_0$

● 対立仮説H_1：$\mu \neq \mu_0$（両側検定）、

$\mu > \mu_0$（右片側検定）、$\mu < \mu_0$（左片側検定）のいずれか※

※与条件などによる

手順2 **有意水準αの設定**

通常は$\alpha=0.05$とする。

手順3 **統計量の計算**

データの平均値$\bar{x}$および母分散σ^2により、Z_0を求める。

$$Z_0=\frac{\bar{x}-\mu_0}{\sqrt{\dfrac{\sigma^2}{n}}}$$

手順4 **検定**

手順３で求めた検定統計量と確率分布表から求めた棄却限界値とを比較して判定する。

● H_1：$\mu \neq \mu_0$の場合、$|Z_0| \geqq Z(\alpha/2)$であれば有意となり、H_0：$\mu=\mu_0$は棄却される。すなわちH_1が成立する。

● H_1：$\mu > \mu_0$の場合、$Z_0 \geqq Z(\alpha)$であれば有意となり、H_0：$\mu=\mu_0$は棄却される。すなわちH_1が成立する。

● H_1：$\mu < \mu_0$の場合、$Z_0 < -Z(\alpha)$であれば有意となり、H_0：$\mu=\mu_0$は棄却される。すなわちH_1が成立する。

手順5 **推定（点推定、区間推定）**

● 点推定　：$\hat{\mu}=\bar{x}$

● 区間推定：$\bar{x}-Z(\alpha/2)\sqrt{\dfrac{\sigma^2}{n}} \leqq \mu \leqq \bar{x}+Z(\alpha/2)\sqrt{\dfrac{\sigma^2}{n}}$

〈例１〉平均値が変わったか否かの検定（母集団の分散が既知の場合）

母平均：$\mu_0=9.0$、母分散：$\sigma^2=1.0$の母集団からデータ数：$n=10$のサンプルをとった結果、標本平均値：$\bar{x}=10.0$であった。ばらつきは変化していないものとする。このとき、平均値が変化したかどうかの検定の手順は以下のとおりである。

手順1 **仮説の設定**

● 帰無仮説H_0：$\mu=\mu_0$

● 対立仮説H_1：$\mu \neq \mu_0$(両側検定)

手順2 **有意水準αの設定**

$\alpha=0.05$とする。

手順3 **統計量の計算**

$$Z_0=\frac{\bar{x}-\mu_0}{\sqrt{\dfrac{\sigma^2}{n}}}=\frac{10.0-9.0}{\sqrt{\dfrac{1.0}{10}}}\fallingdotseq 3.16$$

手順4 検定

$\mu \neq \mu_0$→両側検定　　正規分布表(P.71参照)より、$P=0.025 \rightarrow K_P=$ 1.960となる。$\alpha=P$、$Z(\alpha/2)=K_P$

$Z_0>K_P$となることから、**有意となる**(平均値が変化した)。

帰無仮説は棄却され、対立仮説が成立する。

手順5 推定

- 点推定：$\hat{\mu}=\bar{x}=10.0$
- 区間推定：$\bar{x}-Z(\alpha/2)\sqrt{\dfrac{\sigma^2}{n}} \leqq \mu \leqq \bar{x}+Z(\alpha/2)\sqrt{\dfrac{\sigma^2}{n}}$

(上限)　$\mu_U=\bar{x}+Z(\alpha/2)\sqrt{\dfrac{\sigma^2}{n}}=10.0+1.960\sqrt{\dfrac{1.0}{10}} \fallingdotseq 10.0+0.62=10.62$

(下限)　$\mu_L=\bar{x}-Z(\alpha/2)\sqrt{\dfrac{\sigma^2}{n}}=10.0-1.960\sqrt{\dfrac{1.0}{10}} \fallingdotseq 10.0-0.62=9.38$

よって、$9.38 \leqq \mu \leqq 10.62$

〈例2〉平均値が大きくなったか否かの検定(母集団の分散が既知の場合)

母平均：$\mu_0=9.0$、母分散：$\sigma^2=1.0$の母集団からデータ数：$n=10$のサンプルをとった結果、標本平均値：$\bar{x}=10.0$であった。ばらつきは変化していないものとする。このとき、平均値が大きくなったかどうかの検定の手順は以下のとおりである。

手順1 仮説の設定

- 帰無仮説H_0：$\mu=\mu_0$
- 対立仮説H_1：$\mu>\mu_0$(右片側検定)

手順2 有意水準αの設定

$\alpha=0.05$とする。

手順3 統計量の計算

$$Z_0=\frac{\bar{x}-\mu_0}{\sqrt{\dfrac{\sigma^2}{n}}}=\frac{10.0-9.0}{\sqrt{\dfrac{1.0}{10}}} \fallingdotseq 3.16$$

手順4 検定

$\mu>\mu_0$→片側検定　　正規分布表より、$P=0.05 \rightarrow K_P=1.645$となる。

$\alpha=P$、$Z(\alpha)=K_P$

(参考)正規分布表(一部抜粋)

(II) PからK_Pを求める表

P	0.001	0.005	0.010	0.025	0.050	0.100	0.200	0.300	0.400
K_P	3.090	2.576	2.326	1.960	1.645	1.282	0.842	0.524	0.253

$Z_0>K_P$となることから、**有意となる**(平均値が大きくなった)。

帰無仮説は棄却され、対立仮説が成立する。

手順5 推定

- 点推定：$\hat{\mu}=\bar{x}=10.0$
- 区間推定：$\bar{x}-Z(\alpha/2)\sqrt{\dfrac{\sigma^2}{n}} \leqq \mu \leqq \bar{x}+Z(\alpha/2)\sqrt{\dfrac{\sigma^2}{n}}$

(上限)　$\mu_U=\bar{x}+Z(\alpha/2)\sqrt{\dfrac{\sigma^2}{n}}=10.0+1.645\sqrt{\dfrac{1.0}{10}} \fallingdotseq 10.0+0.52=10.52$

(下限)　$\mu_L=\bar{x}-Z(\alpha/2)\sqrt{\dfrac{\sigma^2}{n}}=10.0-1.645\sqrt{\dfrac{1.0}{10}} \fallingdotseq 10.0-0.52=9.48$

よって、$9.48 \leqq \mu \leqq 10.52$

計量値データに基づく検定と推定

2) 1つの母分散に関する検定と推定

1つの母分散に関する検定と推定は(正規分布に従う際)、$\chi^2=S/\sigma^2$(カイ)が、自由度$\phi(=n-1)$のχ^2分布に従うことを活用して、母分散に関する情報を多く持っている平方和Sおよび分散Vを用いて行います。

手順1 仮説の設定

- 帰無仮説H_0：$\sigma^2=\sigma_0^2$
- 対立仮説H_1：$\sigma^2 \neq \sigma_0^2$(両側検定)、

$\sigma^2>\sigma_0^2$(右片側検定)、$\sigma^2<\sigma_0^2$(左片側検定)のいずれか※

※与条件などによる

手順2 **有意水準αの設定**

通常は$\alpha=0.05$とする。

手順3 **統計量の計算**

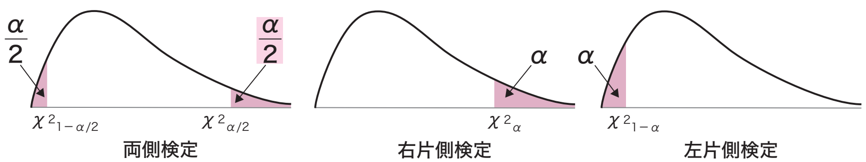
平方和Sを用いて、χ_0^2を求める。

$$\chi_0^2=\frac{S}{\sigma_0^2}$$

手順4 **検定**

手順3で求めた検定統計量と確率分布表から求めた棄却限界値とを比較して判定する。

- H_1：$\sigma^2\neq\sigma_0^2$の場合、$\chi_0^2\geqq\chi^2(n-1,\ \alpha/2)$または$\chi_0^2\leqq\chi^2(n-1,\ 1-\alpha/2)$であれば有意となり（両側検定）、$H_0$：$\sigma^2=\sigma_0^2$は棄却される。すなわち$H_1$が成立する。
- H_1：$\sigma^2>\sigma_0^2$の場合、$\chi_0^2\geqq\chi^2(n-1,\ \alpha)$であれば有意となり（右片側検定）、$H_0$：$\sigma^2=\sigma_0^2$は棄却される。すなわち$H_1$が成立する。
- H_1：$\sigma^2<\sigma_0^2$の場合、$\chi_0^2\leqq\chi^2(n-1,\ 1-\alpha)$であれば有意となり（左片側検定）、$H_0$：$\sigma^2=\sigma_0^2$は棄却される。すなわち$H_1$が成立する。



手順5 **推定（点推定、区間推定）**

- 点推定　：$\hat{\sigma}^2=V=\dfrac{S}{n-1}$　　（$\hat{\sigma}^2$はσ^2の推定値）
- 区間推定：$\chi^2=S/\sigma^2$は、自由度$\phi(=n-1)$のχ^2分布に従うので、

$pr\{\chi^2(n-1,\ 1-\alpha/2)<S/\sigma^2<\chi^2(n-1,\ \alpha/2)\}$

$pr\{S/\chi^2(n-1,\ \alpha/2)<\sigma^2<S/\chi^2(n-1,\ 1-\alpha/2)\}$

となることから、

（上限）　$\sigma^2_U=S/\chi^2(n-1,\ 1-\alpha/2)$

（下限）　$\sigma^2_L=S/\chi^2(n-1,\ \alpha/2)$

〈例〉母分散が大きくなったか否かの検定

ある工程の管理特性は、平均値$\mu=10.00$、分散$\sigma^2=1.00$である。最近、現場から、この特性値のばらつきが大きくなったとの問題提起があったため、サンプリングを行い31個のデータを得たところ、平均値$\bar{x}=10.3$、（不偏）分散$V=1.3$、平方和$S=39$であった。このとき、工程のばらつきが大きくなったか否かの検定の手順は以下のとおりである。

手順1 **仮説の設定**

- 帰無仮説H_0：$\sigma^2=\sigma_0^2$
- 対立仮説H_1：$\sigma^2>\sigma_0^2$（右片側検定）

手順2 **有意水準αの設定**

$\alpha=0.05$とする。

手順3 **統計量の計算**

$\chi_0^2=39/1.00=$**39**

手順4 **検定**

$\sigma^2>\sigma_0^2$→（ばらつきが大きくなったか否かを検定するので）右片側検定

χ^2表において（$\alpha=P$）として、$\chi^2(\phi,\ P)=\chi^2(30,\ 0.05)=$**43.7730**となる。

※データ数$n=31$個なので、自由度$\phi=n-1=30$となる

（参考）χ^2表（一部抜粋）

ϕ \ P	0.995	0.990	0.985	0.975	0.970	0.950	0.050
1	0.00003927	0.0001571	0.0003535	0.0009821	0.001414	0.003932	3.84146
2	0.010025	0.020101	0.030227	0.050636	0.060918	0.102587	5.99146
3	0.071722	0.114832	0.151574	0.215795	0.245795	0.351846	7.81473
14	4.07467	4.66043	5.05724	5.62873	5.85563	6.57063	23.6848
15	4.60092	5.22935	5.65342	6.26214	6.50322	7.26094	24.9958
16	5.14221	5.81221	6.26280	6.90766	7.16251	7.96165	26.2962
17	5.69722	6.40776	6.88415	7.56419	7.83241	8.67176	27.5871
18	6.26480	7.01491	7.51646	8.23075	8.51199	9.39046	28.8693
19	6.84397	7.63273	8.15884	8.90652	9.20044	10.1170	30.1435
20	7.43384	8.26040	8.81050	9.59078	9.89708	10.8508	31.4104
21	8.03365	8.89720	9.47076	10.2829	10.6013	11.5913	32.6706
22	8.64272	9.54249	10.1390	10.9823	11.3125	12.3380	33.9244
23	9.26042	10.1957	10.8147	11.6886	12.0303	13.0905	35.1725
24	9.88623	10.8564	11.4974	12.4012	12.7543	13.8484	36.4150
25	10.5197	11.5240	12.1867	13.1197	13.4840	14.6114	37.6525
26	11.1602	12.1981	12.8821	13.8439	14.2190	15.3792	38.8851
27	11.8076	12.8785	13.5833	14.5734	14.9592	16.1514	40.1133
28	12.4613	13.5647	14.2900	15.3079	15.7042	16.9279	41.3371
29	13.1211	14.2565	15.0019	16.0471	16.4538	17.7084	42.5570
30	13.7867	14.9535	15.7188	16.7908	17.2076	18.4927	43.7730

$\chi_0^2 < \chi^2(30, 0.05) = 43.7730$ となることから、有意ではない（ばらつきは大きくなっていない）。帰無仮説は棄却されず、対立仮説は成立しない。

手順5 推定

- 点推定　：$\hat{\sigma}^2 = V = 1.3$　　（$\hat{\sigma}^2$はσ^2の推定値）
- 区間推定：$S/\chi^2(n-1,\ \alpha/2) < \sigma^2 < S/\chi^2(n-1,\ 1-\alpha/2)$

（上限）　$\sigma^2_U = S/\chi^2(\phi,\ 1-\alpha/2) = 39/\chi^2(30, 0.975) = 39/16.7908 \fallingdotseq 2.32$

（下限）　$\sigma^2_L = S/\chi^2(\phi,\ \alpha/2) = 39/\chi^2(30, 0.025) = 39/46.9792 \fallingdotseq 0.83$

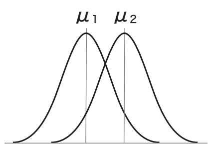
よって、$0.83 < \sigma^2 < 2.32$

計量値データに基づく検定と推定

3）2つの母平均の差に関する検定と推定

２つの集団の母平均の違いを調べたいときに、２つの母平均の差に関する検定と推定を行います。２つの集団は正規分布に従っていて、その平均値は互いに独立していることが前提となります。

※独立していない場合は、データに対応がある場合の検定・推定を行います。



２つの集団それぞれからランダムにサンプルを取り、それぞれの分散V_1、V_2と、平均値$\bar{x}_1$、$\bar{x}_2$を比較します。

$\bar{x}_1$、$\bar{x}_2$は互いに独立しており、$\bar{x}_1 - \bar{x}_2$の分布は正規分布$N(\mu_1 - \mu_2,\ \sigma_1^2/n_1 + \sigma_2^2/n_2)$に従いますので、これを標準化すると、

$$Z_0 = \frac{\bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$$ となります。

σ_1^2、σ_2^2が未知の場合は、t検定を行います。母分散が等しいと考えられる場合と、等しいかどうかわからない場合があるので、２つの場合に分けて検定と推定を行います。

①母分散が等しいと考えられる場合

手順1 仮説の設定

- 帰無仮説H_0：$\mu_1 = \mu_2$
- 対立仮説H_1：$\mu_1 \neq \mu_2$（両側検定）、
$\mu_1 > \mu_2$（右片側検定）、$\mu_1 < \mu_2$（左片側検定）のいずれか※
※与条件などによる

手順2 有意水準αの設定

通常は$\alpha = 0.05$とする。

手順3 統計量等の計算

平均値$\bar{x}_1$、$\bar{x}_2$および平方和S_1、S_2を計算し、共通の分散（併合分散）を計算する。

$$V = s^2 = \frac{S_1 + S_2}{(n_1 - 1) + (n_2 - 1)}$$

次に、検定統計量を計算する。

$$t_0 = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{V\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$$

手順4 検定

H_1：$\mu_1 \neq \mu_2$の場合、$|t_0| > t(\phi_1 + \phi_2,\ \alpha)$であれば有意となり、$H_0$：$\mu_1 = \mu_2$は棄却される。すなわち$H_1$が成立する。

H_1：$\mu_1 < \mu_2$の場合、$t_0 \leqq -t(\phi_1 + \phi_2,\ 2\alpha)$であれば有意となり、$H_0$：$\mu_1 = \mu_2$は棄却される。すなわち$H_1$が成立する。　※$\phi = n - 1$

H_1：$\mu_1 > \mu_2$の場合、$t_0 \geqq t(\phi_1 + \phi_2,\ 2\alpha)$であれば有意となり、$H_0$：$\mu_1 = \mu_2$は棄却される。すなわち$H_1$が成立する。

手順5 推定（点推定、区間推定）

- 点推定　：$\widehat{\mu_1 - \mu_2} = \bar{x}_1 - \bar{x}_2$　　（$\widehat{\mu_1 - \mu_2}$は$\mu_1 - \mu_2$の推定値）
- 区間推定：$(\bar{x}_1 - \bar{x}_2) - t(\phi_1 + \phi_2,\ \alpha)\sqrt{V\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}$
$\leqq \mu_1 - \mu_2$
$\leqq (\bar{x}_1 - \bar{x}_2) + t(\phi_1 + \phi_2,\ \alpha)\sqrt{V\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}$
$\phi = (n_1 - 1) + (n_2 - 1) = \phi_1 + \phi_2$

〈例〉2つの機械A、Bで製造された部品の重量に差があるかどうかを調べることになり、以下のデータを得た。

A：100, 101，95, 102，98, 103

B：105, 108, 109, 100, 106

なお、母分散は等しいと考えられるものとする。

手順1 仮説の設定

- 帰無仮説H_0：$\mu_A = \mu_B$　　※機械A、Bの比較なのでμ_A、μ_Bとする
- 対立仮説H_1：$\mu_A \neq \mu_B$（両側検定）

手順2 有意水準αの設定

$\alpha=0.05$とする。

手順3 統計量等の計算

平均値$\bar{x}_A$、$\bar{x}_B$および平方和S_A、S_Bを計算し、共通の分散（併合分散）を計算する。

$$\bar{x}_A = \frac{100+101+95+102+98+103}{6} \fallingdotseq 99.83$$

$$\bar{x}_B = \frac{105+108+109+100+106}{5} = 105.6$$

$$S_A = \Sigma x_{Ai}^2 - (\Sigma x_{Ai})^2 / n_A$$
$$= 100^2+101^2+95^2+102^2+98^2+103^2$$
$$- \frac{(100+101+95+102+98+103)^2}{6} \fallingdotseq 42.83$$

$$S_B = \Sigma x_{Bi}^2 - (\Sigma x_{Bi})^2 / n_B$$
$$= 105^2+108^2+109^2+100^2+106^2$$
$$- \frac{(105+108+109+100+106)^2}{5} = 49.2$$

$$V = s^2 = \frac{S_A+S_B}{(n_A-1)+(n_B-1)} = \frac{42.83+49.2}{(6-1)+(5-1)} = \frac{92.03}{9} \fallingdotseq 10.23$$

次に、検定統計量を計算する。

$$t_0 = \frac{\bar{x}_A - \bar{x}_B}{\sqrt{V\left(\frac{1}{n_A}+\frac{1}{n_B}\right)}} = \frac{99.83-105.6}{\sqrt{10.23\left(\frac{1}{6}+\frac{1}{5}\right)}} \fallingdotseq \frac{-5.77}{1.94} \fallingdotseq -2.97$$

手順4 検定

$t(\phi_A+\phi_B,\ \alpha) = t((6-1)+(5-1), 0.05)$

$= t(9, 0.05) = 2.262$　　　※$\phi = n-1$

（参考）t表（一部抜粋）

自由度ϕと両側確率Pからtを求める表

ϕ \ P	0.50	0.40	0.30	0.20	0.10	0.05
1	1.000	1.376	1.963	3.078	6.314	12.706
2	0.816	1.061	1.386	1.886	2.920	4.303
3	0.765	0.978	1.250	1.638	2.353	3.182
4	0.741	0.941	1.190	1.533	2.132	2.776
5	0.727	0.920	1.156	1.476	2.015	2.571
6	0.718	0.906	1.134	1.440	1.943	2.447
7	0.711	0.896	1.119	1.415	1.895	2.365
8	0.706	0.889	1.108	1.397	1.860	2.306
9	0.703	0.883	1.100	1.383	1.833	2.262
10	0.700	0.879	1.093	1.372	1.812	2.228

$|t_0| = 2.97 > t(9, 0.05) = 2.262$　であることから、有意となり、H_0：$\mu_A = \mu_B$は棄却される。すなわちH_1が成立する。

手順5 推定（点推定、区間推定）

- 点推定　：$\widehat{\mu_A - \mu_B} = \bar{x}_A - \bar{x}_B = 99.83 - 105.6 = -5.77$
- 区間推定：$(\bar{x}_A - \bar{x}_B) - t(\phi_A+\phi_B,\ \alpha)\sqrt{V(1/n_A+1/n_B)} \leqq \mu_A - \mu_B$
  $\leqq (\bar{x}_A - \bar{x}_B) + t(\phi_A+\phi_B,\ \alpha)\sqrt{V(1/n_A+1/n_B)}$

（上限）　$(\mu_A - \mu_B)_U$

$= (x_A - x_B) + t(\phi_A+\phi_B,\ \alpha)\sqrt{V(1/n_A+1/n_B)}$

$= -5.77 + 2.262 \times 1.94 \fallingdotseq -1.38$

（下限）　$(\mu_A - \mu_B)_L$

$= (\bar{x}_A - \bar{x}_B) - t(\phi_A+\phi_B,\ \alpha)\sqrt{V(1/n_A+1/n_B)}$

$= -5.77 - 2.262 \times 1.94 \fallingdotseq -10.16$

よって、$-10.16 \leqq \mu_A - \mu_B \leqq -1.38$

②母分散が等しいかどうかわからない場合

手順1 仮説の設定

- 帰無仮説H_0：$\mu_1 = \mu_2$
- 対立仮説H_1：$\mu_1 \neq \mu_2$（両側検定）、
  $\mu_1 > \mu_2$（右片側検定）、$\mu_1 < \mu_2$（左片側検定）のいずれか
  ※与条件などによる

手順2 有意水準αの設定

$\alpha = 0.05$とする。

手順3 統計量等の計算

平均値$\bar{x}_1$、$\bar{x}_2$および分散V_1、V_2を計算する。

また、分散V_1、V_2および自由度ϕ_1、ϕ_2を用いて、自由度ϕ^*を求める。

この自由度ϕ^*は等価自由度と呼ばれ、以下の式で求められる。

$$\phi^* = (V_1/n_1 + V_2/n_2)^2 / \{(V_1/n_1)^2/\phi_1 + (V_2/n_2)^2/\phi_2\}$$

この計算式をサタスウェイトの方法という。計算が複雑なため、以下の関係式を使って検算を行うとよい。

(ϕ_1とϕ_2のいずれか小さい方)$< \phi^* < \phi_1 + \phi_2$

※なお、必ずしも整数にならないので、補間を行う

〈補間〉　例えば、$\phi^* = 5.6$となった場合、
$t(\phi^*, 0.05) = t(5.6, 0.05)$の値を、
5.6は５と６の間の数であることから、
$t(5, 0.05) = 2.571$と$t(6, 0.05) = 2.447$の２つの値を用いて、
以下の補間により求める。
$t(5.6, 0.05) = (1 - 0.6) \times t(5, 0.05) + 0.6 \times t(6, 0.05)$
$= 0.4 \times 2.571 + 0.6 \times 2.447 = 2.4966$
※なお、補間とは、数値と数値の間にあるはずの数値を想定すること、割り出すことという意味である

次に、検定統計量を計算する。

$$t_0 = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{V_1}{n_1} + \dfrac{V_2}{n_2}}}$$

（ウェルチの（t）検定とサタスウェイトの方法）

ウェルチの（t）検定は、２つの標本の母平均の差の検定を行う計算の仕方。２つの標本のサンプルサイズ（n）、平均値（$\bar{x}$）、不偏分散（V）を基にして上の式からt_0値を計算し、「２つの標本の母集団の平均値が等しい」という帰無仮説を検定する。

サタスウェイトの方法は、独立した標本の線形結合の有効自由度を近似計算するために使用されるものである。

手順4 検定

H_1：$\mu_1 \neq \mu_2$の場合、$|t_0| > t(\phi^*, \alpha)$であれば有意となり、帰無仮説$H_0$：$\mu_1 = \mu_2$は棄却される。すなわち対立仮説$H_1$が成立する。

H_1：$\mu_1 < \mu_2$の場合、$t_0 \leqq -t(\phi^*, 2\alpha)$であれば有意となり、$H_0$：$\mu_1 = \mu_2$は棄却される。すなわち$H_1$が成立する。

H_1：$\mu_1 > \mu_2$の場合、$t_0 \geqq t(\phi^*, 2\alpha)$であれば有意となり、$H_0$：$\mu_1 = \mu_2$は棄却される。すなわち$H_1$が成立する。

手順5 推定（点推定、区間推定）

- 点推定　：$\widehat{\mu_1 - \mu_2} = \bar{x}_1 - \bar{x}_2$
- 区間推定：$(\bar{x}_1 - \bar{x}_2) - t(\phi^*, \alpha)\sqrt{V_1/n_1 + V_2/n_2}$
  $\leqq \mu_1 - \mu_2$
  $\leqq (\bar{x}_1 - \bar{x}_2) + t(\phi^*, \alpha)\sqrt{V_1/n_1 + V_2/n_2}$

〈例〉２つの機械A、Bで製造された部品の重量に差があるかどうかを調べることになり、以下のデータを得た。

A：100, 101, 95, 102, 98, 103

B：105, 108, 109, 100, 106

なお、母分散は等しいかどうかわからないものとする。

手順1 仮説の設定

- 帰無仮説H_0：$\mu_A = \mu_B$
- 対立仮説H_1：$\mu_A \neq \mu_B$（両側検定）

手順2 有意水準αの設定

$\alpha = 0.05$とする。

手順3 統計量等の計算

平均値$\bar{x}_A$、$\bar{x}_B$および分散V_A、V_Bを計算する。

$$\bar{x}_A = \frac{100+101+95+102+98+103}{6} \fallingdotseq 99.83$$

$$\bar{x}_B = \frac{105+108+109+100+106}{5} = 105.6$$

$$S_A = \Sigma x_{Ai}^2 - (\Sigma x_{Ai})^2 / n_A$$
$$= 100^2+101^2+95^2+102^2+98^2+103^2 - \frac{(100+101+95+102+98+103)^2}{6} \fallingdotseq 42.83$$

$$S_B = \Sigma x_{Bi}^2 - (\Sigma x_{Bi})^2 / n_B$$
$$= 105^2+108^2+109^2+100^2+106^2 - \frac{(105+108+109+100+106)^2}{5} = 49.2$$

$$V_A = \frac{S_A}{n_A - 1} = \frac{42.83}{6-1} = \frac{42.83}{5} \fallingdotseq 8.57$$

$$V_B = \frac{S_B}{n_B - 1} = \frac{49.2}{5-1} = \frac{49.2}{4} = 12.3$$

$$t_0 = \frac{\bar{x}_A - \bar{x}_B}{\sqrt{\frac{V_A}{n_A} + \frac{V_B}{n_B}}} = \frac{99.83 - 105.6}{\sqrt{\frac{8.57}{6} + \frac{12.3}{5}}} \fallingdotseq -2.93$$

分散 V_A、 V_Bおよび自由度 ϕ_A、 ϕ_Bを用いて、**サタスウェイトの方法**により、自由度 ϕ^*を求める。

$$\phi^* = \left(\frac{V_A}{n_A} + \frac{V_B}{n_B}\right)^2 / \left\{\left(\frac{V_A}{n_A}\right)^2 / \phi_A + \left(\frac{V_B}{n_B}\right)^2 / \phi_B\right\}$$
$$= \left(\frac{8.57}{6} + \frac{12.3}{5}\right)^2 / \left\{\left(\frac{8.57}{6}\right)^2 / 5 + \left(\frac{12.3}{5}\right)^2 / 4\right\} \fallingdotseq 7.87$$

$\phi_A = n_A - 1$、$\phi_B = n_B - 1$

手順4 検定

$|t_0| = 2.93$

$t(\phi^*,\ \alpha) = t(7.87, 0.05) \leqq t(7, 0.05) = 2.365$

補間すると、 $t(7.87, 0.05) = (1-0.87) \times t(7, 0.05) + 0.87 \times t(8, 0.05) = 0.13 \times 2.365 + 0.87 \times 2.306 = 2.31367 \fallingdotseq 2.314$

※本番の検定においては、ϕの小数点以下を四捨五入して求めた答えを「＝」で示してもよい

(参考) t 表(一部抜粋)

自由度 ϕ と両側確率 P から t を求める表

ϕ \ P	0.50	0.40	0.30	0.20	0.10	0.05
1	1.000	1.376	1.963	3.078	6.314	12.706
2	0.816	1.061	1.386	1.886	2.920	4.303
3	0.765	0.978	1.250	1.638	2.353	3.182
4	0.741	0.941	1.190	1.533	2.132	2.776
5	0.727	0.920	1.156	1.476	2.015	2.571
6	0.718	0.906	1.134	1.440	1.943	2.447
7	0.711	0.896	1.119	1.415	1.895	2.365
8	0.706	0.889	1.108	1.397	1.860	2.306

$|t_0| = 2.93 > t(7, 0.05) = 2.365$であるので**有意となり**、$H_0$：$\mu_A = \mu_B$は棄却される。すなわち$H_1$が成立する。

手順5 推定(点推定、区間推定)

- 点推定　：$\widehat{\mu_A - \mu_B} = \bar{x}_A - \bar{x}_B = 99.83 - 105.6 = -5.77$
- 区間推定：$(\bar{x}_A - \bar{x}_B) - t(\phi^*,\ \alpha)\sqrt{V_A/n_A + V_B/n_B} \leqq \mu_A - \mu_B \leqq (\bar{x}_A - \bar{x}_B) + t(\phi^*,\ \alpha)\sqrt{V_A/n_A + V_B/n_B}$

(上限)　$(\mu_A - \mu_B)_U$
$$= (\bar{x}_A - \bar{x}_B) + t(\phi^*,\ \alpha)\sqrt{V_A/n_A + V_B/n_B}$$
$$= -5.77 + 2.314\sqrt{\frac{8.57}{6} + \frac{12.3}{5}} = -1.21$$

※補間した数値 2.314を使う

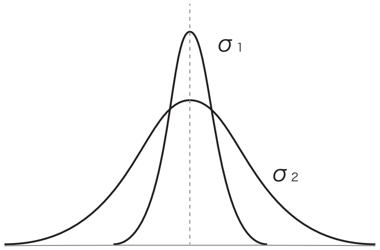
(下限)　$(\mu_A - \mu_B)_L$
$$= (\bar{x}_A - \bar{x}_B) - t(\phi^*,\ \alpha)\sqrt{V_A/n_A + V_B/n_B}$$
$$= -5.77 - 2.314\sqrt{\frac{8.57}{6} + \frac{12.3}{5}} = -10.33$$

よって、$-10.33 \leqq \mu_A - \mu_B \leqq -1.21$

計量値データに基づく検定と推定

4) 2つの母分散の比に関する検定と推定

2つの母分散を比較する場合、母集団のばらつきがほぼ同じであれば問題ありませんが、母集団のばらつきが異なっている場合はこれも踏まえて、2つの母集団の母分散を比較しなければなりません。



それぞれの母集団のサンプルの分散 V_1、V_2の F値、$F=\frac{V_1/\sigma_1^2}{V_2/\sigma_2^2}$が、$F$分布(自由度$\phi_1$(＝$n_1-1$)、$\phi_2$(＝$n_2-1$))に従うことを利用して検定と推定を行います。

手順1 仮説の設定

- 帰無仮説H_0：$\sigma_1^2=\sigma_2^2$
- 対立仮説H_1：$\sigma_1^2\neq\sigma_2^2$（両側検定)、
  $\sigma_1^2>\sigma_2^2$（右片側検定)、$\sigma_1^2<\sigma_2^2$（左片側検定)のいずれか※
  ※与条件などによる

手順2 有意水準αの設定

通常は$\alpha=0.05$とする。

手順3 統計量等の計算

分散V_1、V_2を計算する。V_1、V_2のうち大きい方を分子に持ってくる。

$V_1\geqq V_2$の場合、$F_0=\frac{V_1}{V_2}$($\phi_1=n_1-1$, $\phi_2=n_2-1$)

$V_1<V_2$の場合、$F_0=\frac{V_2}{V_1}$($\phi_1=n_2-1$, $\phi_2=n_1-1$)

手順4 検定

手順3で求めた検定統計量F_0と確率分布表から求めた棄却限界値とを比較して判定する。

- H_1：$\sigma_1^2\neq\sigma_2^2$の場合、$F_0\geqq F(\phi_1, \phi_2; \alpha/2)$であれば有意となり、$H_0$：$\sigma_1^2=\sigma_2^2$は棄却される。すなわち$H_1$が成立する。
- H_1：$\sigma_1^2>\sigma_2^2$の場合、$F_0\geqq F(\phi_1, \phi_2; \alpha)$であれば有意となり、$H_0$：$\sigma_1^2=\sigma_2^2$は棄却される。すなわち$H_1$が成立する。
- H_1：$\sigma_1^2<\sigma_2^2$の場合、$F_0\geqq F(\phi_1, \phi_2; \alpha)$であれば有意となり、$H_0$：$\sigma_1^2=\sigma_2^2$は棄却される。すなわち$H_1$が成立する。

手順5 推定(点推定、区間推定)

- 点推定　：$\hat{\sigma}_1^2/\hat{\sigma}_2^2=V_1/V_2$
- 区間推定：$F=\frac{V_1/\sigma_1^2}{V_2/\sigma_2^2}$は自由度$n_1-1$、$n_2-1$の$F$分布に従うことから、

$$\frac{V_1}{V_2}\times\frac{1}{F(\phi_1, \phi_2; \alpha/2)}\leqq\frac{\sigma_1^2}{\sigma_2^2}\leqq\frac{V_1}{V_2}\times\frac{1}{F(\phi_1, \phi_2; 1-\alpha/2)}$$

$\left(\frac{1}{F(\phi_1, \phi_2; 1-\alpha/2)}=F(\phi_2, \phi_1; \alpha/2)\text{の関係があることを利用して}\right)$

$$\frac{V_1}{V_2}\times\frac{1}{F(\phi_1, \phi_2; \alpha/2)}\leqq\frac{\sigma_1^2}{\sigma_2^2}\leqq\frac{V_1}{V_2}\times F(\phi_2, \phi_1; \alpha/2)$$

〈例〉2つの機械A、Bで製造された部品の重量の分散に差があるかどうかを調べることになり、以下のデータを得た。

A：100, 101, 95, 102, 98, 103

B：105, 108, 109, 100, 106

手順1 仮説の設定

- 帰無仮説H_0：$\sigma_A^2=\sigma_B^2$
- 対立仮説H_1：$\sigma_A^2\neq\sigma_B^2$（両側検定)

手順2 有意水準αの設定

$\alpha=0.05$とする。

手順3 統計量等の計算

平方和S_A、S_Bおよび分散V_A、V_Bを計算する。

$$S_A=\Sigma x_{Ai}^2-(\Sigma x_{Ai})^2/n_A$$
$$=100^2+101^2+95^2+102^2+98^2+103^2-\frac{(100+101+95+102+98+103)^2}{6}\fallingdotseq 42.83$$

$$S_B=\Sigma x_{Bi}^2-(\Sigma x_{Bi})^2/n_B$$
$$=105^2+108^2+109^2+100^2+106^2-\frac{(105+108+109+100+106)^2}{5}=49.2$$

※分散比F_0は、回帰による不偏分散V_Rを残差による不偏分散Veで割った値。$F_0=V_R/Ve$

$$V_A=\frac{S_A}{n_A-1}=\frac{42.83}{6-1}=\frac{42.83}{5}\fallingdotseq 8.57$$

$$V_B=\frac{S_B}{n_B-1}=\frac{49.2}{5-1}=\frac{49.2}{4}=12.3$$

$V_A<V_B$なので、$F_0=V_B/V_A=12.3/8.57\fallingdotseq 1.44$

手順4 検定

手順３で求めた検定統計量と確率分布表から求めた棄却限界値とを比較して判定する。

$F(\phi_A,\ \phi_B;\alpha/2)=F(4,\ 5;0.025)=7.39$

$(\phi_A=n_A-1、\phi_B=n_B-1)$

（参考）*F* 表②（一部抜粋）

$F(\phi_1,\ \phi_2;\alpha)$　　$\alpha=0.025$

ϕ_1＝分子の自由度　　ϕ_2＝分母の自由度

ϕ_2 \ ϕ_1	1	2	3	4	5
1	647.8	799.5	864.2	899.6	921.8
2	38.51	39.00	39.17	39.25	39.30
3	17.44	16.04	15.44	15.10	14.88
4	12.22	10.65	9.98	9.60	9.36
5	10.01	8.43	7.76	7.39	7.15

$F_0=1.44<F(4,\ 5;0.025)=7.39$となるため、有意とならず、$H_0:\sigma_A^2=\sigma_B^2$は棄却されない。

手順5 推定（点推定、区間推定）

（検定の結果が有意ではないが、参考までに行ってみる。）

- 点推定　：$\hat\sigma_A^2/\hat\sigma_B^2=V_A/V_B=\dfrac{8.57}{12.3}\fallingdotseq 0.70$
- 区間推定：$F=\dfrac{V_A/\sigma_A^2}{V_B/\sigma_B^2}$は自由度$n_A-1$、$n_B-1$の$F$分布に従うことから、

$$\frac{V_A}{V_B}\times\frac{1}{F(\phi_A,\ \phi_B;\alpha/2)}\leqq\frac{\sigma_A^2}{\sigma_B^2}\leqq\frac{V_A}{V_B}\times F(\phi_B,\ \phi_A;\alpha/2)$$

$$0.70\times\frac{1}{9.36}\leqq\frac{\sigma_A^2}{\sigma_B^2}\leqq 0.70\times 7.39$$

$$0.07\leqq\frac{\sigma_A^2}{\sigma_B^2}\leqq 5.17$$

5）データに対応がある場合の検定と推定

データに対応がある場合とは、２つの母集団の独立性が成立していない場合です。母平均の差$\mu_d(=\mu_1-\mu_2)$を調べるためのデータが同一対象物に対して組になっていて、互いに関係しているということです。

手順1 仮説の設定

- 帰無仮説H_0：$\mu_d=0$
- 対立仮説H_1：$\mu_d\neq 0$（両側検定）

　$\mu_d>0$（右片側検定）、$\mu_d<0$（左片側検定）のいずれか※

　※与条件などによる

手順2 有意水準αの設定

通常は$\alpha=0.05$とする。

手順3 統計量等の計算

データの差の平均値$\bar d$、分散V_dを計算する。

統計量t_0を計算する。

$t_0=\bar d/\sqrt{V_d/n}$

手順4 検定

手順３で求めた検定統計量t_0と確率分布表から求めた棄却限界値$t(\phi,\alpha)$とを比較して判定する。　※$\phi=n-1$

- H_1：$\mu_d\neq 0$（両側検定）の場合、$|t_0|\geqq t(\phi,\ \alpha)$であれば有意となり、$H_0$：$\mu_d=0$は棄却される。すなわち$H_1$が成立する。
- H_1：$\mu_d>0$（右片側検定）の場合、$t_0\geqq t(\phi,\ 2\alpha)$であれば有意となり、$H_0$：$\mu_d=0$は棄却される。すなわち$H_1$が成立する。
- H_1：$\mu_d<0$（左片側検定）の場合、$t_0\leqq -t(\phi,\ 2\alpha)$であれば有意となり、$H_0$：$\mu_d=0$は棄却される。すなわち$H_1$が成立する。

手順5 推定（点推定、区間推定）

- 点推定　：$\widehat{\mu_1-\mu_2}=\bar d$
- 区間推定：$\bar d-t(\phi,\ \alpha)\sqrt{\dfrac{V_d}{n}}\leqq\mu_1-\mu_2\leqq\bar d+t(\phi,\ \alpha)\sqrt{\dfrac{V_d}{n}}$

〈例〉A工場における部品Bの受入検査において、外注先の検査による合否判定をしていたが、A工場の検査室にて外注先の検査の合否判定の正確性を確認することになった。外注先が検査のための測定を行ったあるロットの部品Bのサンプルを5つ取り、検査室でも測定を行ったところ、測定値(㎝)は下表のとおりであった。

サンプル番号	1	2	3	4	5
外注先	3.5	3.4	3.8	3.9	3.5
検査室	3.2	3.3	3.5	3.7	3.0
データの差	0.3	0.1	0.3	0.2	0.5

外注先と検査室とで測定値の母平均に差があるかどうかを検討する。

手順1 仮説の設定

- 帰無仮説H_0：$\mu_d=0$
- 対立仮説H_1：$\mu_d \neq 0$（差があるかどうかを検定するので）両側検定

手順2 有意水準αの設定

$\alpha=0.05$とする。

手順3 統計量等の計算

データの差の平均値$\bar{d}$を求める。

$$\bar{d}=\frac{0.3+0.1+0.3+0.2+0.5}{5}=\frac{1.4}{5}=0.28$$

$$S=\Sigma d_i^2-(\Sigma d_i)^2/n$$
$$=0.3^2+0.1^2+0.3^2+0.2^2+0.5^2-\frac{(0.3+0.1+0.3+0.2+0.5)^2}{5}$$
$$\fallingdotseq 0.48-0.39=0.09$$

分散$V_d=S/(n-1)=0.09/4=0.0225$

統計量$t_0=\bar{d}/\sqrt{V_d/n}=0.28/\sqrt{0.0225/5}\fallingdotseq 4.17$

手順4 検定

手順3で求めた検定統計量と確率分布表から求めた棄却限界値とを比較して判定する。 ※$\phi=n-1$

- H_1：$\mu_d\neq 0$（両側検定）の場合、$|t_0|\geqq t(\phi,\ \alpha)$であれば有意となり、$H_0$：$\mu_d=0$は棄却される。すなわち$H_1$が成立する。

$t(\phi,\ \alpha)=t(4,0.05)=2.776$

（参考）t表（一部抜粋）

自由度ϕと両側確率Pからtを求める表

ϕ \ P	0.50	0.40	0.30	0.20	0.10	0.05
1	1.000	1.376	1.963	3.078	6.314	12.706
2	0.816	1.061	1.386	1.886	2.920	4.303
3	0.765	0.978	1.250	1.638	2.353	3.182
4	0.741	0.941	1.190	1.533	2.132	2.776
5	0.727	0.920	1.156	1.476	2.015	2.571
6	0.718	0.906	1.134	1.440	1.943	2.447
7	0.711	0.896	1.119	1.415	1.895	2.365
8	0.706	0.889	1.108	1.397	1.860	2.306
9	0.703	0.883	1.100	1.383	1.833	2.262
10	0.700	0.879	1.003	1.372	1.812	2.228

$|t_0|=4.17>t(4,0.05)=2.776$

となるので、有意となり、H_0：$\mu_d=0$は棄却される。

手順5 推定（点推定、区間推定）

- 点推定　：$\widehat{\mu_1-\mu_2}=\bar{d}=0.28$
- 区間推定：$\bar{d}-t(\phi,\ \alpha)\sqrt{\dfrac{V_d}{n}}\leqq\mu_1-\mu_2\leqq\bar{d}+t(\phi,\ \alpha)\sqrt{\dfrac{V_d}{n}}$

（上限）　$\bar{d}+t(\phi,\ \alpha)\sqrt{\dfrac{V_d}{n}}=0.28+2.776\sqrt{\dfrac{0.0225}{5}}\fallingdotseq 0.47$

（下限）　$\bar{d}-t(\phi,\ \alpha)\sqrt{\dfrac{V_d}{n}}=0.28-2.776\sqrt{\dfrac{0.0225}{5}}\fallingdotseq 0.09$

よって、$0.09<\mu_1-\mu_2<0.47$

（2）計数値データに基づく検定と推定

計数値は離散的な分布であり、連続的な分布として表せませんが、連続的な分布に近似して検定と推定を行います。

1）母不適合品率に関する検定と推定

母集団から、試料nを採取して検査した際に、不適合品がx個あったときに、x/nが母不適合品率P_0と等しいかを検定します。

手順1 **仮説の設定**

- 帰無仮説H_0：$P=P_0$
- 対立仮説H_1：$P \neq P_0$（両側検定）、
  $P>P_0$（右片側検定）、$P<P_0$（左片側検定）のいずれか※
  ※与条件などによる

手順2 **有意水準αの設定**

通常は$\alpha=0.05$とする。

手順3 **統計量等の計算**

サンプル中の不適合品数xより、$p=x/n$を計算する。
統計量u_0を計算する。
$u_0=(x-nP_0)/\sqrt{nP_0(1-P_0)}$

手順4 **検定**

手順３で求めた検定統計量と確率分布表から求めた棄却限界値とを比較して判定する。

- H_1：$P \neq P_0$（両側検定）の場合、$|u_0| \geqq u(\alpha/2)$であれば有意となり、H_0：$P=P_0$は棄却される。すなわちH_1が成立する。
- H_1：$P>P_0$（右片側検定）の場合、$u_0 \geqq u(\alpha)$であれば有意となり、H_0：$P=P_0$は棄却される。すなわちH_1が成立する。
- H_1：$P<P_0$（左片側検定）の場合、$u_0 \leqq -u(\alpha)$であれば有意となり、H_0：$P=P_0$は棄却される。すなわちH_1が成立する。

手順5 **推定（点推定、区間推定）**

- 点推定　：$\hat{P}=p=x/n$
- 区間推定：$p-u(\alpha/2)\sqrt{\dfrac{p(1-p)}{n}} \leqq P \leqq p+u(\alpha/2)\sqrt{\dfrac{p(1-p)}{n}}$

〈例〉ある工場でビニル製品の加工不適合品率は$P=0.1$であった。今回、不適合品率の改善のため、ラインの一部を変更して加工を行い、サンプル$n=100$をとって検査したところ、不適合品数$x=15$個であった。不適合品率は変化したといえるかを確認する。

手順1 **仮説の設定**

- 帰無仮説H_0：$P=P_0$
- 対立仮説H_1：$P \neq P_0$

（変化したかどうかなので（増えたか、減ったかではないので））両側検定

手順2 **有意水準αの設定**

$\alpha=0.05$とする。

手順3 **統計量等の計算**

$u_0=(x-nP_0)/\sqrt{nP_0(1-P_0)}$
$=(15-100\times0.1)/\sqrt{100\times0.1(1-0.1)}$
$=5/\sqrt{9} \fallingdotseq$ **1.67**

手順4 **検定**

$P \neq P_0$→両側検定　$|u_0|<u(\alpha/2)=u(0.05/2)=u(0.025)=$**1.960**
となることから**有意ではなく**、帰無仮説は棄却されず、対立仮説は棄却される。

（参考）正規分布表（一部抜粋）
※両側検定なので、$\alpha/2=0.025$となる。

（II）PからK_Pを求める表

P	0.001	0.005	0.010	0.025	0.050	0.100	0.200	0.300	0.400
K_P	3.090	2.576	2.326	1.960	1.645	1.282	0.842	0.524	0.253

手順5 **推定（点推定、区間推定）**

- 点推定　：$\hat{P}=p=x/n=15/100=$**0.15**
- 区間推定：$p-u(\alpha/2)\sqrt{\dfrac{p(1-p)}{n}} \leqq P \leqq p+u(\alpha/2)\sqrt{\dfrac{p(1-p)}{n}}$

（上限）$P_U=p+1.960\sqrt{\dfrac{p(1-p)}{n}}=0.15+1.960\sqrt{\dfrac{0.15(1-0.15)}{100}}$

$=0.15+1.960\sqrt{\dfrac{0.1275}{100}} \fallingdotseq$ **0.22**

（下限）$P_L=p-1.960\sqrt{\dfrac{p(1-p)}{n}}=0.15-1.960\sqrt{\dfrac{0.15(1-0.15)}{100}}$

$=0.15-1.960\sqrt{\dfrac{0.1275}{100}} \fallingdotseq$ **0.08**

よって、**0.08**$\leqq P \leqq$**0.22**

2）2つの母不適合品率の違いに関する検定と推定

２つの母不適合品率(P_1、P_2)から、それぞれn_1、n_2個のサンプルを抜取検査した際に、不適合品がそれぞれx_1、x_2個あったときに、この２つの母不適合品率(P_1、P_2)が等しいかどうかを検定します。

手順１ 仮説の設定

- 帰無仮説H_0：$P_1 = P_2$
- 対立仮説H_1：$P_1 \neq P_2$（両側検定）、
  $P_1 > P_2$（右片側検定）、$P_1 < P_2$（左片側検定）のいずれか※
  ※与条件などによる

手順２ 有意水準αの設定

通常は$\alpha = 0.05$とする。

手順３ 統計量等の計算

サンプル中の不適合品数x_1、x_2より、不適合品率P_1、P_2を計算する。

$\overline{P} = (x_1 + x_2)/(n_1 + n_2)$

統計量u_0を計算する。

$$u_0 = (P_1 - P_2)/\sqrt{\overline{P}(1-\overline{P})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$$

手順４ 検定

手順３で求めた検定統計量と確率分布表から求めた棄却限界値とを比較して判定する。

- H_1：$P_1 \neq P_2$(両側検定)の場合、$|u_0| \geqq u(\alpha/2) = 1.960$であれば有意となり、$H_0$：$P_1 = P_2$は棄却される。すなわち$H_1$が成立する。
- H_1：$P_1 > P_2$(右片側検定)の場合、$u_0 \geqq u(\alpha) = 1.645$ であれば有意となり、H_0：$P_1 = P_2$は棄却される。すなわちH_1が成立する。
- H_1：$P_1 < P_2$(左片側検定)の場合、$u_0 \leqq -u(\alpha) = -1.645$であれば有意となり、$H_0$：$P_1 = P_2$は棄却される。すなわち$H_1$が成立する。

（参考）正規分布表（一部抜粋）

(Ⅱ)PからK_Pを求める表

P	0.001	0.005	0.010	0.025	0.050	0.100	0.200	0.300	0.400
K_P	3.090	2.576	2.326	1.960	1.645	1.282	0.842	0.524	0.253

手順５ 推定（点推定、区間推定）

- 点推定　：$\widehat{P_1 - P_2} = P_1 - P_2 = \frac{x_1}{n_1} - \frac{x_2}{n_2}$
- 区間推定：

$$(P_1 - P_2) - 1.960\sqrt{\frac{P_1(1-P_1)}{n_1}+\frac{P_2(1-P_2)}{n_2}}$$
$$\leqq P_1 - P_2$$
$$\leqq (P_1 - P_2) + 1.960\sqrt{\frac{P_1(1-P_1)}{n_1}+\frac{P_2(1-P_2)}{n_2}}$$

〈例〉２つのラインで生産されている自動車部品がある。各ラインから500個ずつサンプルを抜取検査したところ、Aラインではx_1=12個、Bラインではx_2=10個の不適合品が見つかった。ラインによって母不適合品率に違いがあるかを確認する。

手順１ 仮説の設定

- 帰無仮説H_0：$P_1 = P_2$
- 対立仮説H_1：$P_1 \neq P_2$（両側検定）

手順２ 有意水準αの設定

$\alpha = 0.05$とする。

手順３ 統計量等の計算

不適合品数x_1＝12、x_2＝10より、

不適合品率P_1＝12／500＝0.024、P_2＝10／500＝0.020

$\overline{P} = (x_1 + x_2)/(n_1 + n_2)$

＝(12＋10)／(500＋500)＝22／1000＝0.022

統計量　$u_0=(P_1-P_2)/\sqrt{\bar{P}(1-\bar{P})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$

$=(0.024-0.020)/\sqrt{0.022(1-0.022)\left(\frac{1}{500}+\frac{1}{500}\right)}$

$=0.004/\sqrt{0.022\times0.978\times\frac{2}{500}}$

$\fallingdotseq0.004/\sqrt{0.0000861}\fallingdotseq0.43$

手順4 検定

$P\neq P_0$(両側検定)の場合、$|u_0|<u(\alpha/2)=u(0.05/2)=u(0.025)=1.960$となることから有意ではなく、帰無仮説は棄却されず、対立仮説は棄却される。

(参考)正規分布表(一部抜粋)

※両側検定なので、$\alpha/2=0.025$となる。

(II)PからK_Pを求める表

P	0.001	0.005	0.010	0.025	0.050	0.100	0.200	0.300	0.400
K_P	3.090	2.576	2.326	1.960	1.645	1.282	0.842	0.524	0.253

手順5 推定(点推定、区間推定)

- 点推定　：$\widehat{P_1-P_2}=P_1-P_2=0.024-0.020=0.004$
- 区間推定：$(P_1-P_2)-1.960\sqrt{\frac{P_1(1-P_1)}{n_1}+\frac{P_2(1-P_2)}{n_2}}\leqq P_1-P_2$

$\leqq(P_1-P_2)+1.960\sqrt{\frac{P_1(1-P_1)}{n_1}+\frac{P_2(1-P_2)}{n_2}}$

(上限)　$(P_1-P_2)_U=(P_1-P_2)+1.960\sqrt{\frac{P_1(1-P_1)}{n_1}+\frac{P_2(1-P_2)}{n_2}}$

$=(0.024-0.020)+1.960\sqrt{\frac{0.024(1-0.024)}{500}+\frac{0.020(1-0.020)}{500}}$

$\fallingdotseq0.004+1.960\sqrt{0.000086}\fallingdotseq0.004+0.01818=0.02218$

(下限)　$(P_1-P_2)_L=(P_1-P_2)-1.960\sqrt{\frac{P_1(1-P_1)}{n_1}+\frac{P_2(1-P_2)}{n_2}}$

$=(0.024-0.020)-1.960\sqrt{\frac{0.024(1-0.024)}{500}+\frac{0.020(1-0.020)}{500}}$

$\fallingdotseq0.004-1.960\sqrt{0.000086}\fallingdotseq0.004-0.01818=-0.01418$

よって、$-0.01418\leqq P_1-P_2\leqq0.02218$

計数値データに基づく検定と推定

3）母不適合品数に関する検定と推定

製品の欠損箇所や事故等のサンプル(n個)の不適合数(欠点数)xは母不適合数(母欠点数)をmとするポアソン分布に従います。ポアソン分布は、$m\geqq5$だと正規分布$N(m,\ m)$に近似できることを利用して、検定と推定を行います。

手順1 仮説の設定

- 帰無仮説H_0：$m=m_0$
- 対立仮説H_1：$m\neq m_0$（両側検定)、

$m>m_0$（右片側検定)、$m<m_0$（左片側検定)のいずれか※

※与条件などによる

手順2 有意水準αの設定

通常は$\alpha=0.05$とする。

手順3 統計量等の計算

サンプル中の不適合品数の平均値$\bar{c}$を計算する。$\bar{c}=x/n$

統計量u_0を計算する。

$u_0=(\bar{c}-m_0)/\sqrt{m_0/n}$

手順4 検定

手順３で求めた検定統計量と確率分布表から求めた棄却限界値とを比較して判定する。

- H_1：$m\neq m_0$(両側検定)の場合、$|u_0|\geqq u(\alpha/2)=1.960$であれば有意となり、$H_0$：$m=m_0$は棄却される。すなわち$H_1$が成立する。
- H_1：$m>m_0$(右片側検定)の場合、$u_0\geqq u(\alpha)=1.645$であれば有意となり、H_0：$m=m_0$は棄却される。すなわちH_1が成立する。
- H_1：$m<m_0$(左片側検定)の場合、$u_0\leqq-u(\alpha)=-1.645$であれば有意となり、$H_0$：$m=m_0$は棄却される。すなわち$H_1$が成立する。

手順5 推定(点推定、区間推定)

- 点推定　：$\hat{m} = \bar{c} = x / n$
- 区間推定：$\bar{c} - u(\alpha/2)\sqrt{\frac{\bar{c}}{n}} \leqq m \leqq \bar{c} + u(\alpha/2)\sqrt{\frac{\bar{c}}{n}}$

$$\bar{c} - 1.960\sqrt{\frac{\bar{c}}{n}} \leqq m \leqq \bar{c} + 1.960\sqrt{\frac{\bar{c}}{n}}$$

〈例〉ある工程で加工されるガラス板には、従来1㎡当たり平均5個のキズが発生していた。これを改善し、その効果を確認するために、サンプルを抜き取り、10㎡のガラス板を検査したところ、合計12個のキズが発生していた。母不適合数が減少したかどうかを確認する。

手順1 仮説の設定

- 帰無仮説H_0：$m = m_0$
- 対立仮説H_1：$m < m_0$（減少したかどうかなので）左片側検定

手順2 有意水準αの設定

$\alpha = 0.05$とする。

手順3 統計量等の計算

サンプル中の不適合品数の平均値$\bar{c}$を計算する。

$\bar{c} = x / n = 12/10 = 1.2$

統計量u_0を計算する。

$$u_0 = (\bar{c} - m_0) / \sqrt{m_0 / n} = (1.2 - 5) / \sqrt{\frac{5}{10}} \fallingdotseq -\frac{3.8}{0.71} \fallingdotseq -5.35$$

手順4 検定

$m < m_0$（左片側検定）で、$u_0 \leqq -u(\alpha) = -u(0.05) = -1.645$となるので、有意となり、帰無仮説$H_0$：$m = m_0$は棄却される。すなわち対立仮説$H_1$が成立する。

（参考）正規分布表（一部抜粋）

※左片側なので、符号がマイナス（−）となる。

(II) PからK_Pを求める表

P	0.001	0.005	0.010	0.025	0.050	0.100	0.200	0.300	0.400
K_P	3.090	2.576	2.326	1.960	1.645	1.282	0.842	0.524	0.253

手順5 推定(点推定、区間推定)

- 点推定　：$\hat{m} = \bar{c} = x / n = 1.2$
- 区間推定：$\bar{c} - u(\alpha/2)\sqrt{\frac{\bar{c}}{n}} \leqq m \leqq \bar{c} + u(\alpha/2)\sqrt{\frac{\bar{c}}{n}}$

（上限）　$(m)_U = \bar{c} + 1.960\sqrt{\frac{\bar{c}}{n}} = 1.2 + 1.960\sqrt{\frac{1.2}{10}}$

$\fallingdotseq 1.2 + 0.68 = 1.88$

（下限）　$(m)_L = \bar{c} - 1.960\sqrt{\frac{\bar{c}}{n}} = 1.2 - 1.960\sqrt{\frac{1.2}{10}}$

$\fallingdotseq 1.2 - 0.68 = 0.52$

よって、$0.52 \leqq m \leqq 1.88$

計数値データに基づく検定と推定

4）2つの母不適合数の違いに関する検定と推定

2つの母集団から、それぞれn_1個、n_2個のサンプルを抜取検査した際に、不適合数（欠点数）の合計がそれぞれx_1個、x_2個あったときに、この2つの母不適合数（母欠点数）m_1個、m_2個が等しいかどうかを検定します。

手順1 仮説の設定

- 帰無仮説H_0：$m_1 = m_2$
- 対立仮説H_1：$m_1 \neq m_2$（両側検定）、
  $m_1 > m_2$（右片側検定）、$m_1 < m_2$（左片側検定）のいずれか※
  ※与条件などによる

手順2 有意水準αの設定

通常は$\alpha = 0.05$とする。

手順3 統計量等の計算

不適合数(欠点数) x_1、x_2から、c_1、c_2を計算する。

$\bar{c}=(x_1+x_2)/(n_1+n_2)$

$c_1=x_1/n_1 \qquad c_2=x_2/n_2$

c_1、c_2はサンプルの単位当たり不適合数(欠点数)を示す。

$\bar{c}$はサンプルの単位当たりの全体の不適合数(欠点数)を示す。

統計量u_0を計算する。

$$u_0=(c_1-c_2)/\sqrt{\bar{c}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$$

手順4 検定

手順３で求めた検定統計量と確率分布表から求めた棄却限界値とを比較して判定する。

- H_1：$m_1\neq m_2$(両側検定)の場合、$|u_0|\geqq u(\alpha/2)=1.960$であれば有意となり、$H_0$：$m_1=m_2$は棄却される。すなわち$H_1$が成立する。
- H_1：$m_1>m_2$(右片側検定)の場合、$u_0\geqq u(\alpha)=1.645$ であれば有意となり、H_0：$m_1=m_2$は棄却される。すなわちH_1が成立する。
- H_1：$m_1<m_2$(左片側検定)の場合、$u_0\leqq -u(\alpha)=-1.645$であれば有意となり、$H_0$：$m_1=m_2$は棄却される。すなわち$H_1$が成立する。

手順5 推定

- 点推定　：$\widehat{m_1-m_2}=c_1-c_2=\dfrac{x_1}{n_1}-\dfrac{x_2}{n_2}$
- 区間推定：$(c_1-c_2)-u(\alpha/2)\times\sqrt{\left(\dfrac{c_1}{n_1}+\dfrac{c_2}{n_2}\right)}\leqq m_1-m_2$

$$\leqq(c_1-c_2)+u(\alpha/2)\times\sqrt{\left(\frac{c_1}{n_1}+\frac{c_2}{n_2}\right)}$$

$$(c_1-c_2)-1.960\times\sqrt{\left(\frac{c_1}{n_1}+\frac{c_2}{n_2}\right)}$$

$$\leqq m_1-m_2\leqq(c_1-c_2)+1.960\times\sqrt{\left(\frac{c_1}{n_1}+\frac{c_2}{n_2}\right)}$$

〈例〉ある企業には２つの工場（A工場、B工場）がある。A工場では過去１年間で10件、B工場では過去９か月で20件の災害が発生した。工場によって災害発生件数の違いがあるかどうかを検定する。

手順1 仮説の設定

- 帰無仮説H_0：$m_A=m_B$
- 対立仮説H_1：$m_A\neq m_B$（両側検定）

手順2 有意水準αの設定

$\alpha=0.05$とする。

手順3 統計量等の計算

災害発生件数x_A、x_Bから、c_A、c_B、$\bar{c}$を計算する。

$c_A=10/12\fallingdotseq 0.83 \qquad c_B=20/9\fallingdotseq 2.22$

$\bar{c}=(x_A+x_B)/(n_A+n_B)$

$=(10+20)/(12+9)\fallingdotseq 1.43$

※ n_A、n_Bは月数とする

統計量u_0を計算する。

$$u_0=(c_A-c_B)/\sqrt{\bar{c}\left(\frac{1}{n_A}+\frac{1}{n_B}\right)}$$

$$=(0.83-2.22)/\sqrt{1.43\left(\frac{1}{12}+\frac{1}{9}\right)}\fallingdotseq-\frac{1.39}{0.53}\fallingdotseq-2.62$$

手順4 検定

$m_A\neq m_B$(両側検定)で、$|u_0|>1.960$（正規分布表参照）となるので有意となり、帰無仮説H_0：$m_A=m_B$は棄却され、対立仮説H_1は成立する。

手順5 推定

- 点推定　：$\widehat{m_A-m_B}=c_A-c_B=0.83-2.22=-1.39$
- 区間推定：$(c_A-c_B)-u(\alpha/2)\times\sqrt{\left(\dfrac{c_A}{n_A}+\dfrac{c_B}{n_B}\right)}\leqq m_A-m_B$

$$\leqq(c_A-c_B)+u(\alpha/2)\times\sqrt{\left(\frac{c_A}{n_A}+\frac{c_B}{n_B}\right)}$$

（上限）$(m_A-m_B)_U=(c_A-c_B)+1.960\times\sqrt{\left(\dfrac{c_A}{n_A}+\dfrac{c_B}{n_B}\right)}$

$$=(0.83-2.22)+1.960\times\sqrt{\frac{0.83}{12}+\frac{2.22}{9}}$$

$$\fallingdotseq-1.39+1.960\times0.56=-0.2924$$

$$(\text{下限})\,(m_A - m_B)_L = (c_A - c_B) - 1.960 \times \sqrt{\left(\frac{c_A}{n_A} + \frac{c_B}{n_B}\right)}$$

$$= (0.83 - 2.22) - 1.960 \times \sqrt{\frac{0.83}{12} + \frac{2.22}{9}}$$

$$\fallingdotseq -1.39 - 1.960 \times 0.56 = -3.91$$

よって、$-3.91 \leqq m_A - m_B \leqq -0.2924$

計数値データに基づく検定と推定

5）分割表による検定と推定

製品を適合品と不適合品に分類して、いくつかの母集団での不適合品率を比較したいときには、下のような分割表（２つ以上の変数の関係をまとめた表）を用いて検定と推定を行います。

$l \times m$分割表（例）

列＼行	B_1	……	B_m	計
A_1	x_{11}	……	x_{1m}	$T_{1.}$
：		……		
A_l	x_{l1}	……	x_{lm}	$T_{l.}$
計	$T_{.1}$	……	$T_{.m}$	$T_{..}$

手順１ 仮説の設定

- 帰無仮説H_0：母集団によって差はない（行と列は独立している）。
- 対立仮説H_1：母集団によって差はある（行と列は独立していない）。

手順２ 有意水準αの設定

通常は$\alpha = 0.05$とする。

手順３ 統計量等の計算

期待度数t_{ij}を計算する。　$t_{ij} = T_{i.} \times T_{.j} / T_{..}$

統計量χ_0^2を計算する。

$$\chi_0^2 = \sum_{i=1}^{l} \sum_{j=1}^{m} \frac{(x_{ij} - t_{ij})^2}{t_{ij}}$$

手順４ 検定

$\chi_0^2 \geqq \chi^2(\phi,\ \alpha)$であれば有意となり、$H_0$は棄却される。すなわち$H_1$が成立する。自由度$\phi$＝（行数－１）×（列数－１）

〈例〉２台の機器Ａ、Ｂで部品を製作したところ、適合品と不適合品が下表のとおり発生した。機器Ａ、Ｂによって適合品と不適合品の出方に違いがあるかどうかを検討する。

	適合品	不適合品	合計
A	160	40	200
B	120	20	140
合計	280	60	340

手順１ 仮説の設定

- 帰無仮説H_0：母集団によって差はない（行と列は独立している）。
- 対立仮説H_1：母集団によって差はある（行と列は独立していない）。

手順２ 有意水準αの設定

$\alpha = 0.05$とする。

手順３ 統計量等の計算

期待度数t_{ij}を計算する。　$t_{ij} = T_{i.} \times T_{.j} / T_{..}$

※期待度数とは、行要素の合計や列要素の合計の比率から逆算して期待される度数をいう

２×２分割表

	適合品	不適合品	合計
A	200×280／340 ≒165	200×60／340 ≒35	200
B	140×280／340 ≒115	140×60／340 ≒25	140
合計	280	60	340

正解　（１）ア　（２）ウ　（３）カ　（４）キ　（５）ス
（６）コ　（７）エ　（８）ソ

【問３】 Ｃ研究所の装置Ｄの改善を行った。改善前において、300個中、52個の不適合品が確認された。改善後において、460個中、75個の不適合品が確認された。このことを踏まえ、改善の効果の有無を確認することになった。$\boxed{}$内に入るもっとも適切なものを下の選択肢からひとつ選べ。 ※（１）（２）（８）には符号が入る

手順１ 仮説の設定

改善前の不適合品率をP_1、改善後の不適合品率をP_2とすると、

- 帰無仮説H_0：$P_1 \boxed{(1)} P_2$
- 対立仮説H_1：$P_1 \boxed{(2)} P_2$

手順２ 有意水準αの設定

$\alpha=0.05$とする。

手順３ 統計量等の計算

サンプル中（n_1、n_2）の不適合品数x_1、x_2より、不適合品率P_1、P_2を計算すると、

$P_1=\boxed{(3)}$、$P_2=\boxed{(4)}$

$$\overline{P}=\frac{x_1+x_2}{n_1+n_2}=\boxed{(5)}$$

統計量u_0の計算式は、

$u_0=\boxed{(6)}$

手順４ 検定

$u_0=\boxed{(7)}$、$u_0 \boxed{(8)} u(0.05)=1.645$となることから、$\boxed{(9)}$。

〈選択肢〉

ア．＝　イ．≠　ウ．＜　エ．＞　オ．0.163
カ．0.167　キ．0.173　ク．0.361
ケ．$(x-nP_0)/\sqrt{nP_0(1-P_0)}$
コ．$(P_1-P_2)/\sqrt{\overline{P}(1-\overline{P})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$
サ．有意となり、H_0は棄却される
シ．有意とならず、H_0は棄却されない

正解　（１）ア　（２）エ　（３）キ　（４）オ　（５）カ
（６）コ　（７）ク　（８）ウ　（９）シ

解答・解説

【問１】 仮説の考え方（帰無仮説と対立仮説）

（１）ア　（２）エ　（３）イ　（４）ウ
※（１）と（２）、（３）と（４）はそれぞれ順不同

帰無仮説H_0が正しいのに、棄却してしまう誤りを［ア．第１種の誤り］または［エ．あわてものの誤り］という。一方、対立仮説H_1が正しいのに、棄却してしまう誤りを［イ．第２種の誤り］または［ウ．ぼんやりものの誤り］という。

【問２】 １つの母平均に関する検定と推定

（１）ア　（２）ウ　（３）カ　（４）キ　（５）ス
（６）コ　（７）エ　（８）ソ

帰無仮説は改善の効果がない（この問題では「低下していない」）とするものなので、（１）には「ア．＝」が入る。一方、対立仮説は改善の効果がある（低下している）とするものなので、（２）には「ウ．＜」が入る。

問題文にあるデータ{1.3, 1.4, 1.2, 1.4, 1.6, 1.3, 1.2, 1.3, 1.4, 1.5}の平均を求めると、$\overline{x}=$カ．1.36となる。

また、平方和 $S\left(=\Sigma x_i^2-\dfrac{(\Sigma x_i)^2}{n}\right)$ を求めると、$18.64-\dfrac{13.6^2}{10}=0.144$ となる。

これを用いて分散 $V\left(=\dfrac{S}{n-1}\right)$ を求めると、キ. 0.016となる。

平方和の計算が若干手間ではあるが、実際の問題では平方和の値が与えられる場合も多い。

$$t_0=\text{ス.}\ \frac{\bar{x}-\mu_0}{\sqrt{V/n}}=\frac{1.36-1.4}{\sqrt{\dfrac{0.016}{10}}}=\frac{-0.04}{\sqrt{0.0016}}=\frac{-0.04}{0.04}=\text{コ.}\ -1.00$$

$-t(\phi,\ 2\alpha)=-t(9,\ 0.1)=-1.833$

t_0 エ. $>-t(\phi,\ 2\alpha)$ となるので、ソ. 有意とならず、帰無仮説 H_0 は棄却されない。

【問３】2つの母不適合品率の違いに関する検定と推定

（1）ア　（2）エ　（3）キ　（4）オ　（5）カ
（6）コ　（7）ク　（8）ウ　（9）シ

帰無仮説は改善の効果がないとするものなので、（1）には「ア. ＝」が入る。一方、対立仮説は改善の効果がある（改善前の不適合品率 P_1 ＞改善後の不適合品率 P_2）とするものなので、（2）には「エ. ＞」が入る。

問題文より、

$$P_1=(3)=\frac{52}{300}\fallingdotseq\text{キ.}\ 0.173,\quad P_2=(4)=\frac{75}{460}\fallingdotseq\text{オ.}\ 0.163$$

$$\bar{P}=\frac{x_1+x_2}{n_1+n_2}=(5)=\frac{52+75}{300+460}\fallingdotseq\text{カ.}\ 0.167$$

$$u_0=(6)=\text{コ.}\ (P_1-P_2)/\sqrt{\bar{P}(1-\bar{P})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$$

$$=\frac{0.173-0.163}{\sqrt{0.167(1-0.167)\left(\dfrac{1}{300}+\dfrac{1}{460}\right)}}\fallingdotseq\text{ク.}\ 0.361$$

$u_0=0.361$ ウ. ＜1.645となることから、シ. 有意とならず、H_0 は棄却されない。

I 手法編

第5章 相関分析と単回帰分析

合格のポイント

➡ 相関分析（相関係数の計算、無相関の検定、寄与率による検討）、回帰式の基礎的な計算と分散分析表の作成手順の理解

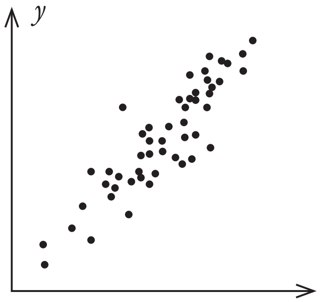
1 相関分析

相関分析とは、**2つの変数(特性値：原因と結果)の関係性の強さを分析する**ことです。2つの変数x(要因変数)、y(結果変数)において、xの連続的な変化に対して、yも連続的に変化する関係が成立する場合、xとyとの間に「**相関がある**」といいます。

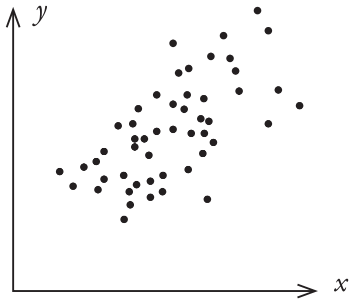
相関係数とは、2つの変数の間に、直線的な関係がどの程度あるかを示す数値で、rで表されます。**$-1 \leqq r \leqq 1$**の範囲をとり、$|r|$が大きいほど相関関係が強いことを示し、$|r|$が小さいほど相関関係が弱いことを示します。

図表5.1 相関係数と相関の強さ

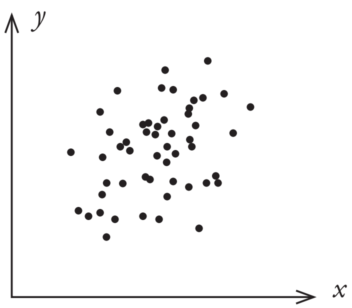
① $r \geqq 0.8$：強い相関がある



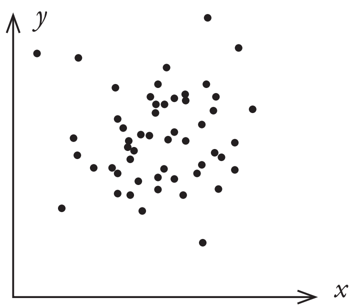
②$0.8 > r \geqq 0.6$：相関がある



③$0.6 > r \geqq 0.4$：弱い相関がある



④ $r < 0.4$：ほとんど相関はない



n個のデータ(x_1, y_1), (x_2, y_2), …, (x_n, y_n)において、「xとyの偏差積和S_{xy}」を「xの偏差平方和の平方根$\sqrt{S_{xx}}$とyの偏差平方和の平方根$\sqrt{S_{yy}}$の積」で割った値を、xとyの**相関係数r**といいます。

$$r=\frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}$$

$$S_{xx}=\Sigma x_i^{\,2}-\frac{(\Sigma x_i)^2}{n}$$

$$S_{yy}=\Sigma y_i^{\,2}-\frac{(\Sigma y_i)^2}{n}$$

$$S_{xy}=\Sigma x_i y_i-\frac{\Sigma x_i \Sigma y_i}{n}$$

〈例〉xの平均＝5.0、平方和＝4.0、yの平均=6.0、平方和＝9.0、xとyの(偏差)積和＝5.0のときの相関係数は、

相関係数 $$r=\frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}=\frac{5.0}{\sqrt{4.0\times 9.0}}=\frac{5.0}{6.0}\fallingdotseq 0.83$$

また、相関係数rを2乗したものを**寄与率R^2**といいます。寄与率R^2は、0～1の範囲をとり、**要因変数xが結果変数yに及ぼす影響の大きさ**を表します。

相関分析を行うデータにおいて、xは$N(\mu_x, \sigma_x^{\,2})$に従い、yは$N(\mu_y, \sigma_y^{\,2})$に従うと考えます。また母集団におけるxとyの関連度合いを表す母数とし、下式のとおり、母相関係数ρ(ロー)を考えます。

$$\rho=\frac{Cov\,(x,y)}{\sqrt{V(x)\,V(y)}}=\frac{E\left[(x-\mu_x)(y-\mu_y)\right]}{\sqrt{\sigma x^2 \sigma y^2}}$$

分子は**共分散**を示し、分母はxとyの**標準偏差**を示します。共分散をxとyの標準偏差で割ることで**標準化**します。

ρもrと同様、$-1 \leqq \rho \leqq 1$となり、1に近づくほど母集団としてxとyの正の相関が強く、-1に近づくほど負の相関が強く、0に近づくほど相関が弱いことを示します。

※個々のデータから平均値を引いた値を偏差といい、xとyの偏差をかけ合わせた偏差積$(x_i-\bar{x})(y_i-\bar{y})$の総和を$x$と$y$の偏差積和といい、$S_{xy}$で表す

1）母相関係数ρに関するt分布に基づく無相関検定

手順1 仮説の設定

- 帰無仮説H_0：$\rho = 0$（無相関）
- 対立仮説H_1：$\rho \neq 0$（相関関係がある）

手順2 有意水準αの設定

通常は$\alpha = 0.05$とする。

手順3 統計量等の計算

n個のデータ$(x_1,\ y_1)$，$(x_2,\ y_2)$，…，$(x_n,\ y_n)$における相関係数rを計算する。

$$r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}$$

統計量t_0を計算する。

$$t_0 = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

手順4 検定

$|t_0| \geqq t(\phi,\ \alpha)$であれば有意となり、$H_0$は棄却される。すなわち$H_1$が成立する。

2）グラフによる相関の検定

対になる２つの変数、要因(x)、特性(y)のグラフがある場合は、これを利用して相関の検定を行うことができます。

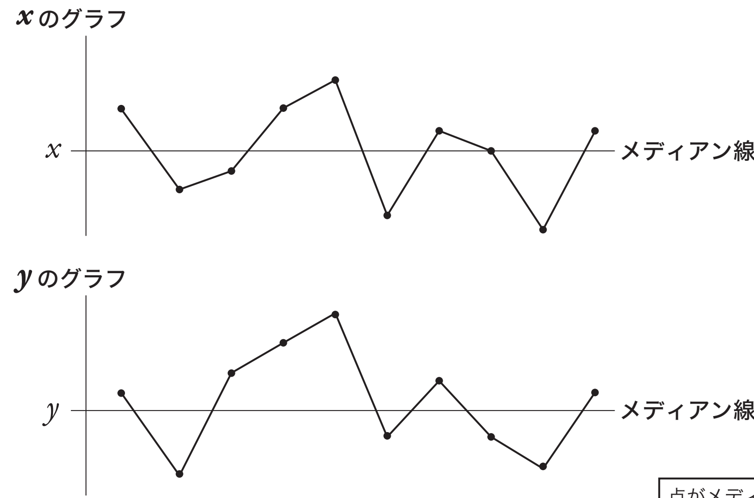
グラフによる相関の検定方法には、xとyの中央値（メディアン）を利用する**大波の相関**と、前のデータからの増減を利用する**小波の相関**があります。

①大波の相関の検定方法

手順1 中央線（メディアン線）の作成

下図のxとyを対にしたグラフにおいて、点の数を上下に２等分する中央線（メディアン線）を引きます（図表５.２参照）。

図表5.2 大波の相関



x	＋	−	−	＋	＋	−	＋	0	−	＋
y	＋	−	＋	＋	＋	−	＋	−	−	＋
$x \times y$	＋	＋	−	＋	＋	＋	＋	0	＋	＋

点がメディアン線の上にある場合は＋を、下にある場合はマイナスをつけて、対になる点の符号どうしをかけ合わせる（$x \times y$）。

手順2 符号の積の系列の作成等

中央線の上側にある点に＋、下側にある点に−を付けて、その符号の積の系列を作ります（図表5.２下の「$x \times y$」表参照）。次にxとyの積の「＋」の数n_+と「−」の数n_-を数えます。中央線上に点があれば０として、これは数えないこととします。この例によると、$n_+ = 8$、$n_- = 1$となります。

手順3 判定

符号検定表と比較して判定を行います。符号検定表とは、統計的に相関があるか否かを判定する表です。合計$N(= n_+ + n_-)$に該当する行から、判定数n_sを読み取ります。

明変数x、目的変数y)を線(目的変数yについて説明変数xを使った式)で表すことです。

説明変数xが1つの場合を**単回帰分析**といい、説明変数が2つ以上ある場合を**重回帰分析**といいます。QC検定®2級では、**単回帰分析**が出題対象となっています。

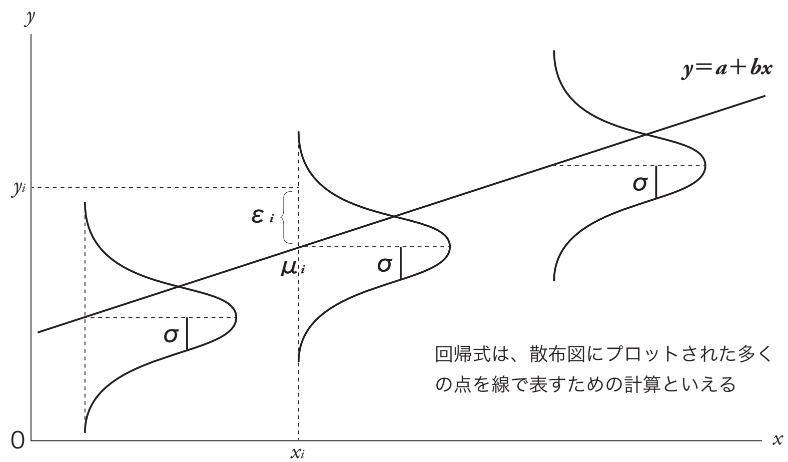
単回帰式は、$y = a + b\,x + \varepsilon$(エプシロン)と表されます。
(a:切片、b:傾き、ε:残差)

※回帰式に含まれる係数a、bを回帰係数という

単回帰分析では、すべてのデータの残差εが小さくなるように、aとbを算出します。データの数がn個ある場合、i番目の値を(x_i, y_i)とすると、真の回帰式から得られる値は、(x_i, $a + b\,x_i$)となります。これらを用いると、残差ε_iは、

ε_i=実際のデータの値−真の回帰式から得られる値$= y_i - (a + b\,x_i)$

と表すことができます。



なお、上図は単回帰モデルと呼ばれ、yの母平均は直線$y = a + b\,x$上にあり、$x = x_i$と指定すると、母平均$y_i = a + b\,x_i$が定まり、それに正規分布$N(0, \sigma^2)$に従う残差ε_iが加わってデータy_iが得られることになります。

すべてのデータの残差εを小さくするために、各データの二乗和を考え、この二乗和が最小となるようなaとbを算出します。この方法を**最小二乗法**といいます。

aとbを求めると、

$a = y$の平均値$- b \times (x$の平均値)

$b = S_{xy} / S_{xx}$

となります。

上述の回帰式$y = a + b\,x + \varepsilon$を変形して、$y_i - \bar{y} = b\,(x_i - \bar{x}) + \varepsilon$と表すこともできます。これにより、$x_i$における$y_i$の母平均の推定値$\hat{\mu}_i$は以下の式で表されます。

$\hat{\mu}_i = \bar{y} + b\,(x_i - \bar{x})$

実測値は回帰線上にはないことから、総変動S_Tは以下の通り分解されます。

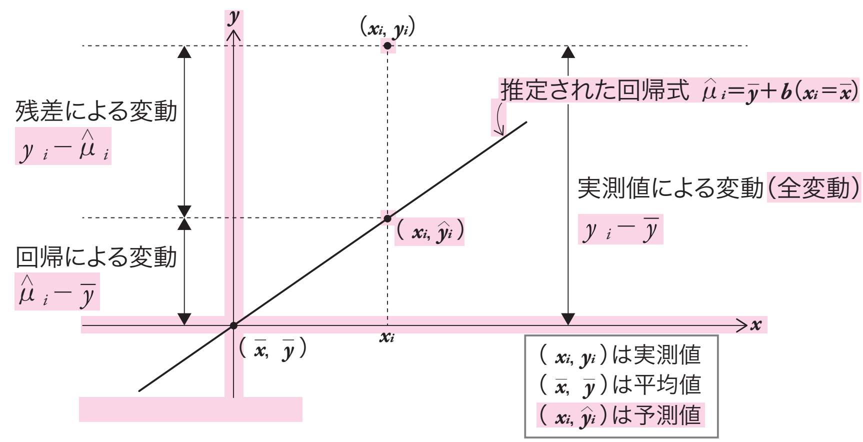
$S_T = \Sigma\,(y_i - \bar{y})^2$ → 実測値の変動を示します。

$S_R = \Sigma\,(\hat{\mu}_i - \bar{y})^2$ → 回帰による変動を示します。

$S_e = \Sigma\,(y_i - \hat{\mu}_i)^2$ → 残差による変動を示します。

上記の3つの変動において、$S_T = S_R + S_e$の関係が成立します。

図表5.5 変動の分解



$$S_R = \Sigma\,(\hat{\mu}_i - \bar{y})^2 = \Sigma\,\{\bar{y} + b\,(x_i - \bar{x}) - \bar{y}\}^2 = b^2 S_{xx}$$

$b = \dfrac{S_{xy}}{S_{xx}}$なので、$S_R = \dfrac{(S_{xy})^2}{S_{xx}}$

$$S_e = S_T - S_R = S_{yy} - \frac{(S_{xy})^2}{S_{xx}}$$

となり、yの総変動が回帰による変動と残差による変動とに分解されたことを示しています。

残差 ε については、

- 不偏性　$(E(\varepsilon_i)=0)$
- 等分散性 $(V(\varepsilon_i)=\sigma^2)$
- 独立性　$(Cov(\varepsilon_i,\ \varepsilon_j)=0,\ i\neq j)$
- 正規性　$(\varepsilon_i \sim N(0,\ \sigma^2))$　※正規分布に従う。「$\sim N(0,\ \sigma^2)$」とは、「$N(0,\ \sigma^2)$に従う」という意味

を満たす必要があります。

2）分散分析（単回帰分析）の手順

分散分析（単回帰分析）は、説明変数 x と目的変数（実測値）y の関係を、グラフにする際に、直線的な関係が予想される場合に回帰に意味があるかどうかの検定と、検定の結果、意味があった場合に回帰係数の推定を行うものです。

データ表（例）

標本	説明変数 x	目的変数（実測値）y
1	x_1	y_1
2	x_2	y_2
:	:	:
n	x_n	y_n

手順1　各平方和の計算

総変動　　　：$S_T = S_{yy}$

回帰による変動：$S_R = (S_{xy})^2 / S_{xx}$

残差による変動：$S_e = S_T - S_R = S_{yy} - \dfrac{(S_{xy})^2}{S_{xx}}$

手順2　各自由度の計算

全体の自由度　　：ϕ_T＝総データ数－1＝$n-1$

回帰による自由度：$\phi_R = 1$

残差による自由度：$\phi_e = n-2$

手順3　各（不偏）分散 V と分散比 F_0 の計算

$$V_R = \frac{S_R}{\phi_R} = S_R$$

$$V_e = \frac{S_e}{\phi_e} = \frac{S_e}{n-2}$$

$$F_0 = \frac{V_R}{V_e}$$

手順4　分散分析表の作成

手順１～３の結果を下記のような分散分析表にまとめ、検定と推定を行う。

要因	平方和	自由度	平均平方	分散比
回帰	S_R	1	$V_R = S_R$	$F_0 = \dfrac{V_R}{V_e}$
残差	S_e	$n-2$	$V_e = \dfrac{S_e}{n-2}$	
計	S_T	$n-1$	−	

手順5　検定

分散比 F_0 と F 表の $F(1,\ n-2;\alpha)$ を比較します。

$F_0 > F(1,\ n-2;\alpha)$ であれば、回帰による変動が残差による変動よりも全変動に与える影響が大きいといえ、回帰曲線は予測に役立つ＝回帰による変動が有意である（意味のあること）と判定します。

手順6　推定

回帰による変動が有意である（意味のあること）と判定した場合、回帰係数の推定を行います。

回帰係数の推定値 a，b $(y = a + bx)$ は、

$b = S_{xy} / S_{xx}$

$a = y$ の平均値 $- b \times (x$ の平均値$)$

により求めます。

〈例〉10組のデータにおいて、x の平均4.0、平方和6.0、y の平均5.0、平方和7.0、x と y の偏差積和が6.0のとき、分散分析表を作成して、回帰に意味があるのか否かの検定と、検定の結果、意味があった場合には回帰係数の推定を行う。

手順1 **各平方和の計算**

総変動　　　　：$S_T = S_{yy} = 7.0$

回帰による変動：$S_R = (S_{xy})^2 / S_{xx} = (6.0)^2 / 6.0 = 6.0$

残差による変動：$S_e = S_T - S_R = S_{yy} - (S_{xy})^2 / S_{xx}$

$= 7.0 - 6.0 = 1.0$

手順2 **各自由度の計算**

全体の自由度　：$\phi_T =$ 総データ数 $-1 = n - 1 = 10 - 1 = 9$

回帰による自由度：$\phi_R = 1$

残差による自由度：$\phi_e = n - 2 = 10 - 2 = 8$

手順3 **各(不偏)分散Vと分散比F_0の計算**

$$V_R = \frac{S_R}{\phi_R} = S_R = 6.0$$

$$V_e = \frac{S_e}{\phi_e} = \frac{S_e}{n-2} = \frac{1.0}{10-2} = \frac{1}{8}$$

$$F_0 = \frac{V_R}{V_e} = \frac{6.0}{1/8} = 48$$

手順4 **分散分析表の作成**

要因	平方和	自由度	平均平方	分散比
回帰	$S_R = 6.0$	1	$V_R = 6.0$	$F_0 = \frac{V_R}{V_e} = 48$
残差	$S_e = 1.0$	$n - 2 = 8$	$V_e = \frac{1}{8}$	
計	$S_T = 7.0$	$n - 1 = 9$	−	

手順5 **検定**

分散比F_0とF表の$F(1, n-2; \alpha)$を比較する。

$F(1, n-2; \alpha) = F(1, 8; 0.05) = 5.32$

$F_0 = 48 > F(1, 8; 0.05) = 5.32$となることから、回帰による変動が有意である(意味のあること)と判定する。

手順6 **推定**

回帰による変動が有意である(意味のあること)と判定されたので、回帰係数の推定を行う。

回帰係数の推定値a, b($y = a + bx$)は、

$b = S_{xy} / S_{xx} = 6.0 / 6.0 = 1$

$a = y$の平均値$- b \times$(xの平均値)$= 5.0 - 1.0 \times 4.0 = 1$

よって、$y = 1 + x$となる。

なお、単回帰分析では、説明変数xと目的変数yに比例関係があり、データ範囲外の領域でも比例関係が継続すると仮定している。

例えば、$x = 20.0$を$y = 1 + x$に代入すると、$y = 21.0$となるが、xおよびyの平均を大きく外れてしまう。このように、回帰式を導いたデータの範囲外の数値を代入することを「外挿」というが、誤った結果を含んでいる可能性があることから信憑性に問題があるため、できるだけ避けるべきである。

3）寄与率および残差の検討

回帰分析で得られた回帰式がどの程度の精度であるかを評価する指標として、寄与率R^2があります。

$R^2 = S_R / S_T$

回帰による変動(S_R)を総変動(S_T)で割ったもので、総変動における回帰による変動の割合を示し、0から1の間の値をとります。値が大きいほど回帰式に意味があり、値が小さいほど回帰式に意味がないということになります。なお、**相関係数rの二乗に一致します($R^2 = r^2$)**。

また、残差の大きさについても検討する必要があります。残差とは**目的変数(実測値)yの値と回帰式によって予測したyの値との差**を示します。

なお、残差については次の3つを行います。

- 残差が正規分布に従っているかどうかの検討
- 残差と説明変数xが関係ないかどうかの検討
- 残差の時間的変化に「クセ」があるかどうかの検討

※残差は、測定値から推定値を引いた値

練習問題

赤シートで正解を隠して設問に答えてください(解説はP.119から)。

【問１】　ある工程の要因 x と品質特性 y に関する x と y の対のデータを10組抽出した。これにより以下の統計量を得た。
$\Sigma x=55$、$\Sigma y=47$、$\Sigma x^2=349$、$\Sigma y^2=257$、$\Sigma xy=289$
このとき、[　　]内に入る数値を答えよ。

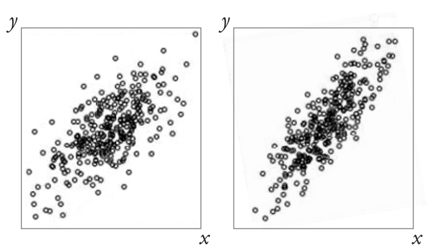
①要因 x の偏差平方和 S_{xx}＝[（１）]、品質特性 y の偏差平方和 S_{yy}＝[（２）]、x と y の偏差積和 S_{xy}＝[（３）]となる。これにより、x と y の相関係数 r＝[（４）]となる。
② x を説明変数、y を目的変数として、回帰式 $y=a+bx$ を推定すると、y＝[（５）]＋[（６）]× x 　回帰式の寄与率は[（７）]となる。
③②で推定された回帰式における分散分析表は下表のとおりとなる。

要因	平方和 S	自由度 ϕ	不偏分散 V	分散比 F_0
回帰	S_R＝[（8）]	$\phi_R=1$	V_R＝[（12）]	F_0＝[（14）]
残差	S_e＝[（9）]	ϕ_e＝[（11）]	V_e＝[（13）]	
計	S_T＝[（10）]	$\phi_T=9$	−	

正解　（１）46.5　（２）36.1　（３）30.5　（４）0.74　（５）1.07
（６）0.66　（７）0.55　（８）20.0　（９）16.1　（10）36.1
（11）8　（12）20.0　（13）2.01　（14）9.95

【問２】　次の文章において、[　　]内に入るもっとも適切なものを下の選択肢からひとつ選べ。

右の散布図（A）、（B）において、特性値 x と特性値 y の相関係数 r の絶対値は（A）よりも（B）が[　　]。



（A）　（B）

〈選択肢〉　ア．大きい　イ．小さい

正解　ア

【問３】　次の文章において、[　　]内に入るもっとも適切なものを下の選択肢からひとつ選べ。ただし、（３）には符号が入る。

ある工程の要因 x と品質特性 y に関する x と y の対のデータを５組抽出した。これにより以下の統計量を得た。
$\Sigma x=9$、$\Sigma y=20$、$\Sigma x^2=19$、$\Sigma y^2=84$、$\Sigma xy=37$
要因 x と品質特性 y に関する相関係数 r を求めると、r＝[（１）]となる。次に t 検定を用いて無相関の検定を行う。t 値を計算すると、

$$t_0=\frac{r\sqrt{n-2}}{\sqrt{1-r^2}}=\boxed{（２）}$$

となり、$|t_0|$[（３）]$t(\phi,\ \alpha)\langle\phi=n-2、n=5\rangle$であることから、相関関係が[（４）]といえる。
〈選択肢〉
ア．0.299　イ．0.543　ウ．0.602　エ．1.305　オ．1.746
カ．＞　キ．＜　ク．＝　ケ．ある　コ．ない

正解　（１）ア　（２）イ　（３）キ　（４）コ

解答・解説

【問１】問Ⅰ〜3　相関分析、単回帰分析
（１）46.5　（２）36.1　（３）30.5　（４）0.74　（５）1.07
（６）0.66　（７）0.55　（８）20.0　（９）16.1　（10）36.1
（11）8　（12）20.0　（13）2.01　（14）9.95

$S_{xx}=349-\frac{55^2}{10}=46.5$、$S_{yy}=257-\frac{47^2}{10}=36.1$、$S_{xy}=289-\frac{55\times47}{10}=30.5$、

$$r=\frac{30.5}{\sqrt{46.5\times36.1}}\fallingdotseq0.74$$

②については、①で求めた偏差積和を用いて、a、b をそれぞれ求める。寄与率は相関係数 r を２乗して求める。

【問２】　ア　散布図（A）よりも（B）のほうが傾きが大きく、直線に近いので、相関係数が大きいことは視覚的にわかる。

【問３】　（１）ア　（２）イ　（３）キ　（４）コ　計算式はP.120参照。
$t(\phi,\ \alpha)=t(3, 0.05)=3.182$

計算式の確認

	計算式
偏差平方和 S	$S_{xx}=\Sigma x_i^2-\frac{(\Sigma x_i)^2}{n}$ $S_{yy}=\Sigma y_i^2-\frac{(\Sigma y_i)^2}{n}$ $S_{xy}=\Sigma x_i y_i-\frac{\Sigma x_i \Sigma y_i}{n}$
相関係数 r	$r=\frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}$
寄与率 $R^2(=r^2)$	相関係数 r を２乗する $R^2(=r^2)=\left(\frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}\right)^2=\frac{S_R}{S_{yy}}=\frac{S_R}{S_T}$
統計量 t_0	$t_0=\frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$
単回帰式 $y=a+bx+\varepsilon$ における a：切片、 b：傾き(回帰係数)	$b=\frac{S_{xy}}{S_{xx}}$ $a=y$ の平均値$-b\times(x$ の平均値)
総平方和	$S_T=S_{yy}$
回帰平方和	$S_R=\frac{(S_{xy})^2}{S_{xx}}$
残差平方和	$S_e=S_T-S_R=S_{yy}-\frac{(S_{xy})^2}{S_{xx}}$
自由度 ϕ	全体の自由度：ϕ_T＝総データ数－１＝$n-1$ 回帰による自由度：ϕ_R＝１(＝回帰式の説明変数の数) 残差による自由度：$\phi_e=n-2$
平均平方 V ※	回帰の平均平方：$V_R=\frac{S_R}{\phi_R}=S_R$ 残差の平均平方：$V_e=\frac{S_e}{\phi_e}=\frac{S_e}{n-2}$
分散比	$F_0=\frac{V_R}{V_e}$

※(不偏)分散と平均平方は同義であるが、過去問の出題実績に則り、第Ⅰ章(データの取り方とまとめ方)では「(不偏)分散」と表記し、第５章(単回帰分析)においては「平均平方」と表記する

Ⅰ 手法編

第6章 実験計画法

合格のポイント

- ➡ 取得したデータからの分散分析表の作成、最適条件を決定する計算方法、および F 表の確認の理解
- ➡ 一元、二元配置実験の違い、二元配置実験におけるプーリング前後のデータの構造式と分散分析表から最適組合せの推定までの手順の理解

1 実験計画法の概要

実験計画法とは、効率的で客観的な結論が得られるように、実験を計画する方法です。具体的には、**目標とする特性値に対して影響のありそうな因子をいくつか取り上げて、その主効果や交互作用効果を検定および推定する**ための統計的方法です。少ない実験回数で重要な情報を得るために、実験の計画(実験配置)と実験データの解析(分散分析)を行います。

イギリスの統計学者のロナルド・エイルマー・フィッシャーがこの実験計画法を確立させました。

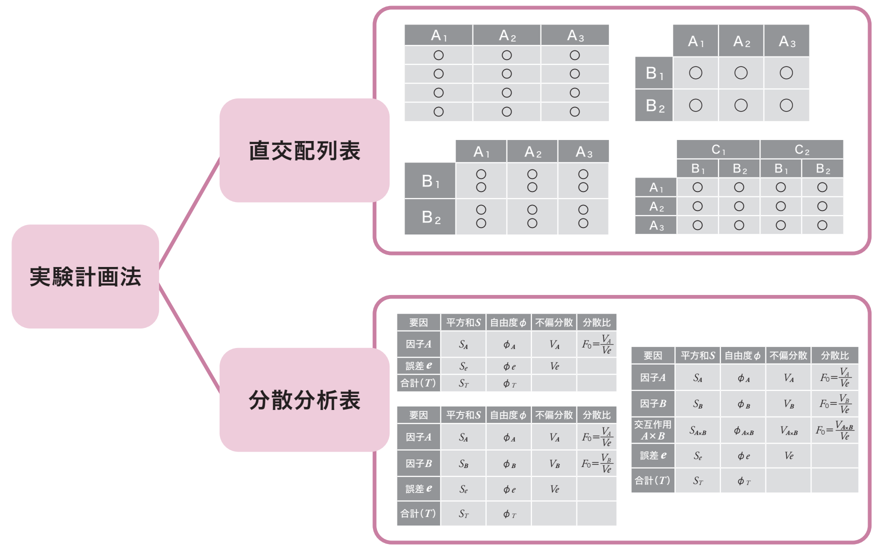
端的にいうと、実験で得られたデータの変化・特徴が、「要因によるものなのか？」「誤差によるものなのか？」を見分ける方法といえます。

図表6.1 実験計画法で見るもの

要因	●実験結果に影響を及ぼす可能性のあるもの(誤差等を含む)
因子	●実験の目的のために取り上げた要因
水準	●因子を量的または質的に変化させた段階で、とくに代表値として選んだ値 ●水準内のばらつきを群内変動、水準間のばらつきを群間変動という
効果	●応答の平均に対する因子の影響 ●(単一因子による)主効果や(ある因子の効果が他の因子の水準に依存する)交互作用がある

実験計画法の確立当時は、計算機が未発達で、全て手計算で行うしかなかったため、重宝されましたが、現在では計算機の発達により、実務で実験計画法が用いられることはほとんどないようです。

実験計画法は、直交配列表と分散分析表の２つの項目で構成されます。



実験の精度を高めるために、**フィッシャーの３原則**を用います(図表6.2参照)。

図表6.2 フィッシャーの３原則の概要

反復の原則	無作為化の原則	局所管理の原則
●同じ条件の実験を繰り返す。 ●実験で繰り返すことで、データ数が増大し、誤差分散が小さくなる。 ●系統誤差※1と偶然誤差※2を判断できる。	●ランダムに行う。 ●誤差要因を一定にすることで、誤差の一定性を確保する(系統誤差を偶然誤差に転化する)。	●できるだけ同じ対象を用いて行う。 ●得たい要因以外の要因が全て均一となる。 ●系統誤差を小さくできる。

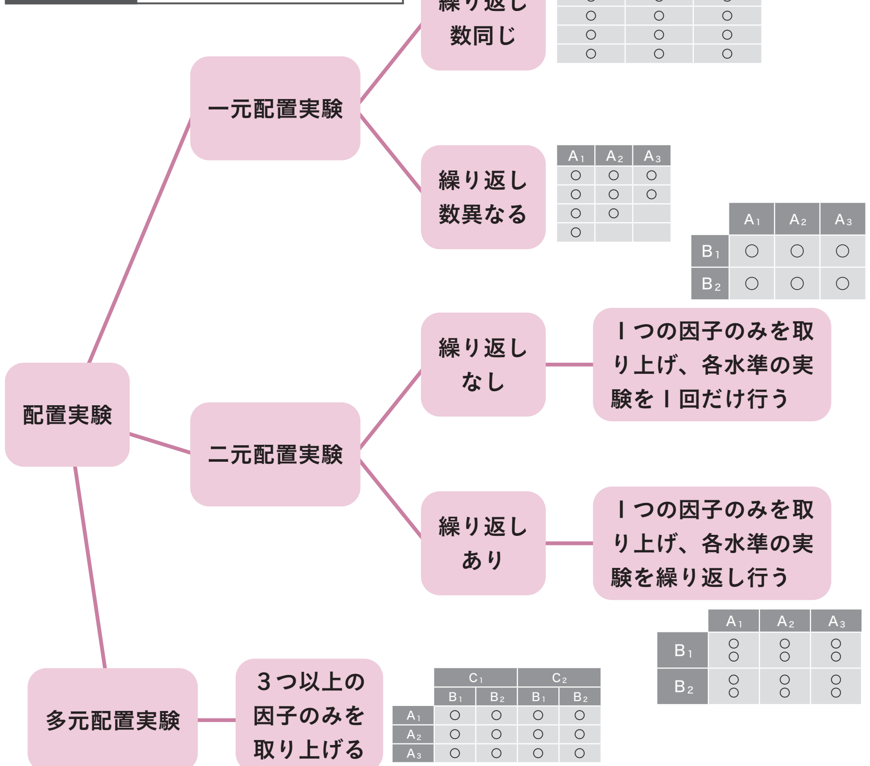
※１系統誤差：処理の違いによる差
※２偶然誤差：たまたま生じる誤差

2 要因配置実験

1）種類

要因配置実験は、配置実験法ともいわれ、図表６．３のように分類されます。ＱＣ検定®２級では、**一元配置実験**および**二元配置実験**(繰り返しなし、繰り返しあり)が試験範囲となります。問題文を読んで、どのパターンに該当するかを適切に判断することが重要です。

図表6．3 要因配置実験一覧(例)



要因配置実験の手順(一例)は、図表６．４のとおりです。

図表6．4 要因配置実験の手順(一例)

①分散分析	②推定
手順１ データの二乗表の作成(二元配置実験(繰り返し有り)の場合は、二元表などの作成)	**手順１** 母平均の点推定
手順２ 修正項(CT)の計算	**手順２** 母平均の区間推定
手順３ 分散分析表の作成(平方和、自由度、平均平方、分散比の計算)	
手順４ 分散分析結果の判定	

データの構造式

	構造式
一元配置実験	$x_{ij}=\mu+\alpha_i+\varepsilon_{ij}$
二元配置実験(繰り返しなし)	$x_{ij}=\mu+\alpha_i+\beta_j+\varepsilon_{ij}$
二元配置実験(繰り返しあり)	$x_{ijk}=\mu+\alpha_i+\beta_j+(\alpha\beta)_{ij}+\varepsilon_{ijk}$※

データ$x_{ij(k)}$は、**平均μ、主効果α_iとβ_j、交互作用$(\alpha\beta)_{ij}$、誤差$\varepsilon_{ij(k)}$**の一次式で表されます。

主効果とは因子単独の効果、**交互作用**とは複数の因子の組み合わせによる効果、**誤差**とはデータから各効果を取り除いた残り物のことです。

2）分散分析表の考え方

実験にあたり、目的とする特性値に影響を与える変動要因の中から、その実験に取り上げた要因を**因子**といい、その要因を質的、量的に変動させる条件を**水準**といいます。

データのばらつきには、一般的に、

- 因子の水準を変えたために生じるばらつき
- 実験を繰り返したときに生じるばらつき

が混在しています。

※下付きの記号ijkは、それぞれi：Aの第i水準、j：Bの第j水準、k：k回目(繰り返しがある場合)を表している。P.126、129、132、136のデータ構造式参照

実験全体のデータが持っているばらつきを**総変動**といい、**総変動**は群間変動と群内移動に分類することができます。

- 群間変動（級間平方和）S_A：水準による応答の平均のばらつき
- 群内変動（誤差平方和）S_e：水準を一定に保った場合の、個々のデータのばらつき

3）一元配置実験（繰り返しの数が同じ）

因子を一つ選び、数通りの水準を設定し、各水準においてランダムに繰り返しの実験を行います。

水準数３、実験の繰り返し数３とすると、A_iにおけるj個目のデータx_{ij}は、次の構造式で観測されると考え、データ表は以下のとおりとなります。

データ＝総平均＋処理の効果＋誤差

$$x_{ij} = \mu + \alpha_i + \varepsilon_{ij}$$

データ表

繰り返し＼水準	A_1	A_2	A_3
1	x_{11}	x_{21}	x_{31}
2	x_{12}	x_{22}	x_{32}
3	x_{13}	x_{23}	x_{33}

要因	平方和S	自由度ϕ	平均平方V	分散比F_0
因子A	$S_A=\Sigma\dfrac{(A_i\text{のデータの合計})^2}{A_i\text{のデータ数}}-CT$	ϕ_A＝水準数－１	$V_A=\dfrac{S_A}{\phi_A}$	$F_0=\dfrac{V_A}{V_e}$
誤差e	$S_e=S_T-S_A$	$\phi_e=\phi_T-\phi_A$	$V_e=\dfrac{S_e}{\phi_e}$	
合計	$S_T=\Sigma$（データの二乗）$-CT$	ϕ_T＝総データ数－１		

CT（修正項）＝（データの合計）^{2}／データ数

〈例〉因子Aを３水準設定し、３回の実験を行ったときの分散分析と推定

データ表

A_1	A_2	A_3
5	5	9
6	4	7
7	6	8

①分散分析

手順１ データの合計表およびデータの二乗表の作成

データの合計表

繰り返し＼水準	A_1	A_2	A_3	総計
1	5	5	9	19
2	6	4	7	17
3	7	6	8	21
合計	18	15	24	57

データの二乗表

繰り返し＼水準	A_1	A_2	A_3	総計
1	25	25	81	131
2	36	16	49	101
3	49	36	64	149
合計	110	77	194	381

手順２ 修正項（CT）の計算

データの合計＝５＋６＋７＋５＋４＋６＋９＋７＋８＝**57**

CT＝（データの合計）^{2}／データ数＝$\dfrac{57^2}{9}$＝**361**

手順３ 分散分析表の作成（平方和、自由度、平均平方、分散比の計算）

まず、各平方和を計算する。

$S_T=\Sigma$（データの二乗）$-CT=381-361=$**20**

$$S_A=\Sigma\frac{(A_i\text{のデータの合計})^2}{A_i\text{のデータ数}}-CT$$

$$=\frac{18^2}{3}+\frac{15^2}{3}+\frac{24^2}{3}-361=375-361=14$$

$S_e=S_T-S_A=20-14=6$

次に、各自由度を計算する。

ϕ_T＝総データ数－１＝９－１＝8

ϕ_A＝水準数－１＝３－１＝2

$\phi_e=\phi_T-\phi_A=8-2=6$

続いて、平均平方と分散比を計算する。

$V_A=S_A/\phi_A=14/2=7$

$V_e=S_e/\phi_e=6/6=1$

これにより、分散比は、$F_0=V_A/V_e=7/1=7$

よって、分散分析表は下表のとおりとなる。

要因	平方和	自由度	平均平方	分散比
因子A	14	2	7	7
誤差e	6	6	1	
合計	20	8		

手順４ 分散分析結果の判定

手順３で得た分散比$F_0=7$とF表の$F(\phi_A,\ \phi_e;\alpha)=F(2,6;0.05)=5.14$を比較すると、$F_0>F(2,6;0.05)$となるので、有意な差があると判定できる。

②推定

手順１ 点推定

①分散分析の結果、因子Aは有意となったので、各水準の母平均μを信頼度95％で推定する。母平均＝各水準の平均値であることから、

A_1水準の母平均＝18／３＝6

A_2水準の母平均＝15／３＝5

A_3水準の母平均＝24／３＝8

手順２ 区間推定

各水準の母平均μの信頼区間幅を信頼度95％で表すと以下の式となる。

$$\widehat{\mu_{Ai}}-t(\phi_e,0.05)\sqrt{\frac{V_e}{n}}<\mu_{Ai}<\widehat{\mu_{Ai}}+t(\phi_e,0.05)\sqrt{\frac{V_e}{n}}$$

（$\widehat{\mu_{Ai}}$は点推定を示す。n：各水準の繰り返し数）

$t(\phi_e,0.05)\sqrt{V_e/n}$において、$n=3$と①手順３で得た、$\phi_e=6$、$V_e=1$を代入すると、$t(6,0.05)\sqrt{1/3}\fallingdotseq2.447\times0.58\fallingdotseq1.42$となる。

$6-1.42<\mu_{A_1}<6+1.42\rightarrow4.58<\mu_{A_1}<7.42$

$5-1.42<\mu_{A_2}<5+1.42\rightarrow3.58<\mu_{A_2}<6.42$

$8-1.42<\mu_{A_3}<8+1.42\rightarrow6.58<\mu_{A_3}<9.42$

4）一元配置実験（繰り返しの数が異なる）

一元配置実験において繰り返しの数が異なる場合を考えてみます。基本的には繰り返しの数が同じ場合と同様の手順となります。

水準数３、実験の繰り返し数３とすると、A_iにおけるj個目のデータx_{ij}は、次の構造式で観測されると考え、データ表は以下のとおりとなります。

データ＝総平均＋処理の効果＋誤差

$$x_{ij}=\mu+\alpha_i+\varepsilon_{ij}$$

データ表

繰り返し＼水準	A_1	A_2	A_3
1	x_{11}	x_{21}	x_{31}
2	x_{12}	x_{22}	x_{32}
3	x_{13}	x_{23}	

要因	平方和S	自由度ϕ	平均平方V	分散比F_0
因子A	$S_A=\Sigma\frac{(A_i\text{のデータの合計})^2}{A_i\text{のデータ数}}-CT$	ϕ_A＝水準数－１	$V_A=\frac{S_A}{\phi_A}$	$F_0=\frac{V_A}{V_e}$
誤差e	$S_e=S_T-S_A$	$\phi_e=\phi_T-\phi_A$	$V_e=\frac{S_e}{\phi_e}$	
合計	$S_T=\Sigma$（データの二乗）$-CT$	ϕ_T＝総データ数－１		

CT（修正項）＝（データの合計）^{2}／データ数

〈例〉因子Aを３水準設定し、３回の実験を行ったときの分散分析と推定

データ表

A_1	A_2	A_3
5	5	9
6	4	8
7	6	

①分散分析

手順１ データの合計表およびデータの二乗表の作成

データの合計表

繰り返し＼水準	A_1	A_2	A_3	総計
1	5	5	9	19
2	6	4	8	18
3	7	6		13
合計	18	15	17	50

データの二乗表

繰り返し＼水準	A_1	A_2	A_3	総計
1	25	25	81	131
2	36	16	64	116
3	49	36		85
合計	110	77	145	332

手順２ 修正項（CT）の計算

データの合計＝５＋６＋７＋５＋４＋６+９＋８＝**50**

CT＝（データの合計数）^{2}／データ数＝50^2／８＝**312.5**

手順３ 分散分析表の作成（平方和、自由度、平均平方、分散比の計算）

まず、各平方和を計算する。

S_T＝Σ（データの二乗）－CT＝332－312.5＝**19.5**

$$S_A=\Sigma\frac{(A_i\text{のデータの合計})^2}{A_i\text{のデータ数}}-CT$$

$$=\frac{18^2}{3}+\frac{15^2}{3}+\frac{17^2}{2}-312.5=327.5-312.5=\mathbf{15}$$

S_e＝S_T－S_A＝19.5－15＝**4.5**

次に、各自由度を計算する。

ϕ_T＝総データ数－１＝８－１＝**7**

ϕ_A＝水準数－１＝３－１＝**2**

ϕ_e＝ϕ_T－ϕ_A＝７－２＝**5**

続いて、平均平方と分散比を計算する。

V_A＝S_A／ϕ_A＝15／２＝**7.5**

V_e＝S_e／ϕ_e＝4.5／５＝**0.9**

これにより、分散比は、F_0＝V_A／V_e＝7.5／0.9≒**8.33**

よって、分散分析表は下表のとおりとなる。

要因	平方和	自由度	平均平方	分散比
因子A	15	2	7.5	8.33
誤差e	4.5	5	0.9	
合計	19.5	7		

手順４ 分散分析結果の判定

手順３で得た分散比F_0=**8.33**とF表の$F(\phi_A, \phi_e; \alpha)=F(2, 5; 0.05)$＝**5.79**を比較すると、$F_0 > F(2, 5; 0.05)$となるので、**有意な差がある**と判定できる。有意水準をα（＝0.05）とする。

②推定

手順１ 点推定

①分散分析の結果、因子Aは有意となったので、各水準の母平均μを信頼度95％で推定する。母平均＝各水準の平均値であることから、

A_1水準の母平均＝18／３＝６

A_2水準の母平均＝15／３＝５

A_3水準の母平均＝17／２＝8.5　　※各水準の合計を繰り返し数で割る

手順2 区間推定

各水準の母平均μの信頼区間幅を信頼度95%で表すと以下の式となる。

$$\widehat{\mu_{Ai}} - t(\phi_e, 0.05)\sqrt{\frac{V_e}{n_i}} < \mu_{Ai} < \widehat{\mu_{Ai}} + t(\phi_e, 0.05)\sqrt{\frac{V_e}{n_i}}$$

（$\widehat{\mu_{Ai}}$は点推定を示す。n_i：各水準の繰り返し数）

$t(\phi_e, 0.05)\sqrt{V_e/n_i}$において、$n_i = 3$と①手順３で得た、$\phi_e = 5$、$V_e = 0.9$を代入すると、$t(5, 0.05)\sqrt{0.9/3} \fallingdotseq 2.571 \times 0.55 \fallingdotseq 1.41$となる。

水準A_1、A_2において、

$6 - 1.41 < \mu_{A1} < 6 + 1.41 \rightarrow 4.59 < \mu_{A1} < 7.41$

$5 - 1.41 < \mu_{A2} < 5 + 1.41 \rightarrow 3.59 < \mu_{A2} < 6.41$

水準A_3において、$n = 2$と①手順３で得た$\phi_e = 5$、$V_e = 0.9$を代入すると、$t(5, 0.05)\sqrt{0.9/2} \fallingdotseq 2.571 \times 0.671 \fallingdotseq 1.72$となる。

$8.5 - 1.72 < \mu_{A3} < 8.5 + 1.72 \rightarrow 6.78 < \mu_{A3} < 10.22$

5）二元配置実験（繰り返しなし）

２つの因子A、Bについて、それぞれa個の水準、b個の水準を選び、全部で$a \times b$個の組み合わせの実験をランダムに行います。

水準数$a = 4$、$b = 3$とすると、A_iの第i水準、B_jの第j水準を組み合わせた水準の下で行った実験のデータx_{ij}は、次の構造式で観測されると考え、データ表は以下のとおりとなります。

データ＝総平均＋処理の効果＋誤差

$$x_{ij} \; = \; \mu \; + \; \alpha_i + \beta_j \; + \varepsilon_{ij}$$

データ表

因子A ＼ 因子B	B_1	B_2	B_3
A_1	x_{11}	x_{21}	x_{31}
A_2	x_{12}	x_{22}	x_{32}
A_3	x_{13}	x_{23}	x_{33}
A_4	x_{14}	x_{24}	x_{34}

２つの因子A、Bの各水準を組み合わせた条件ごとに、１回ずつの実験を行ってデータをとる場合、原則としてランダムな順序で実験を行います。

この実験で得られたデータを分散分析するには、総変動を要因変動（因子Aによる変動、因子Bによる変動）に分離して、各変動の大きさを比較することになります。

要因	平方和S	自由度ϕ	平均平方V	分散比F_0
因子A	$S_A = \Sigma\dfrac{(A_i\text{のデータの合計})^2}{A_i\text{のデータ数}} - CT$	ϕ_A＝水準数－１	$V_A = \dfrac{S_A}{\phi_A}$	$F_0 = \dfrac{V_A}{V_e}$
因子B	$S_B = \Sigma\dfrac{(B_j\text{のデータの合計})^2}{B_j\text{のデータ数}} - CT$	ϕ_B＝水準数－１	$V_B = \dfrac{S_B}{\phi_B}$	$F_0 = \dfrac{V_B}{V_e}$
誤差e	$S_e = S_T - S_A - S_B$	$\phi_e = \phi_T - \phi_A - \phi_B$	$V_e = \dfrac{S_e}{\phi_e}$	
合計	$S_T = \Sigma$（データの二乗）$- CT$	ϕ_T＝総データ数－１		

データ表（因子Aを２水準、因子Bを４水準の実験）

	B_1	B_2	B_3	B_4
A_1	6	10	14	8
A_2	4	8	9	5

①分散分析

手順1 データの合計表およびデータの二乗表の作成

データの合計表

因子A ＼ 因子B	B_1	B_2	B_3	B_4	総計
A_1	6	10	14	8	38
A_2	4	8	9	5	26
合計	10	18	23	13	64

データの二乗表

因子A ＼ 因子B	B_1	B_2	B_3	B_4	総計
A_1	36	100	196	64	396
A_2	16	64	81	25	186
合計	52	164	277	89	582

手順2 修正項（CT）の計算

データの合計＝６＋４＋10＋８＋14＋９＋８＋５＝64

CT＝（データの合計）^{2}／データ数＝64^2／８＝512

手順3 分散分析表の作成（平方和、自由度、平均平方、分散比の計算）

まず、各平方和を計算します。

$S_T=\Sigma$（データの二乗）$-CT=582-512=70$

$$S_A=\Sigma\frac{(A_i\text{のデータの合計})^2}{A_i\text{のデータ数}}-CT$$

$$=\frac{38^2+26^2}{4}-512=\frac{2120}{4}-512=18$$

$$S_B=\Sigma\frac{(B_j\text{のデータの合計})^2}{B_j\text{のデータ数}}-CT$$

$$=\frac{10^2+18^2+23^2+13^2}{2}-512=\frac{1122}{2}-512=49$$

$S_e=S_T-S_A-S_B=70-18-49=3$

次に、各自由度を計算します。

ϕ_T＝総データ数$-1=8-1=7$

ϕ_A＝水準数$-1=2-1=1$

ϕ_B＝水準数$-1=4-1=3$

$\phi_e=\phi_T-\phi_A-\phi_B=7-1-3=3$

続いて、平均平方と分散比を計算します。

$V_A=S_A／\phi_A=18／1=18$

$V_B=S_B／\phi_B=49／3≒16.3$

$V_e=S_e／\phi_e=3／3=1$

これにより、分散比は、

A：$F_0=V_A／V_e=18／1=18$

B：$F_0=V_B／V_e=16.3／1=16.3$

よって、分散分析表は下表のとおりとなります。

要因	平方和	自由度	平均平方	分散比
因子A	18	1	18	18
因子B	49	3	16.3	16.3
誤差e	3	3	1	
合計	70	7		

手順4 分散分析結果の判定

手順３で得た分散比A：$F_0=18$、B：$F_0=16.3$とF表の$F(\phi_A,\ \phi_e;\alpha)=F(1,3;0.05)=10.1$、$F(\phi_B,\ \phi_e;\alpha)=F(3,3;0.05)=9.28$を比較すると、

A：$F_0=18>F(1,3;0.05)=10.1$

B：$F_0=16.3>F(3,3;0.05)=9.28$

となるので、有意な差があると判定できます。

②推定

手順1 最適な組み合わせ条件の選定

特性値が大きい方がよいので、データの合計表において最大値となるA_1B_3を選定します。

手順2 点推定

①分散分析の結果、因子Aは有意となったので、各水準の母平均μを信頼度95％で推定します。母平均＝各水準の平均値であることから、

A_1水準の母平均＝38／４＝9.5　　A_2水準の母平均＝26／４＝6.5

B_1水準の母平均＝10／２＝5　　B_2水準の母平均＝18／２＝9

B_3水準の母平均＝23／２＝11.5　　B_4水準の母平均＝13／２＝6.5

手順3 区間推定

各水準の母平均μの信頼区間幅を信頼度95％で表すと以下の式となります。

$$\widehat{\mu_{Ai}}-t(\phi_e,0.05)\sqrt{\frac{V_e}{n_i}}<\mu_{Ai}<\widehat{\mu_{Ai}}+t(\phi_e,0.05)\sqrt{\frac{V_e}{n_i}}$$

$$\widehat{\mu_{Bj}}-t(\phi_e,0.05)\sqrt{\frac{V_e}{n_j}}<\mu_{Bj}<\widehat{\mu_{Bj}}+t(\phi_e,0.05)\sqrt{\frac{V_e}{n_j}}$$

（$\widehat{\mu_{Ai}}$、$\widehat{\mu_{Bj}}$は点推定を示す。n_i、n_j：各水準の繰り返し数）

Aの場合、$(\phi_e,0.05)\sqrt{V_e/n_i}$において、①分散分析の手順３で得た、$\phi_e=3$、$V_e=1$、$n_i=2$を代入すると、

$$t(3,0.05)\sqrt{\frac{1}{2}}≒3.182\times0.707≒2.250$$　となる。

Bの場合、$(\phi_e,0.05)\sqrt{V_e/n_j}$において、①分散分析の手順３で得た、$\phi_e=3$、$V_e=1$、$n_j=4$を代入すると、

$$t(3,0.05)\sqrt{\frac{1}{4}}=3.182\times0.5=1.591$$　となる。

● A_1水準の母平均の区間推定

9.5−2.250＜μ_{A1}＜9.5＋2.250　　7.25＜μ_{A1}＜11.75

● A_2水準の母平均の区間推定

6.5−2.250＜μ_{A2}＜6.5＋2.250　　4.25＜μ_{A2}＜8.75

● B_1水準の母平均の区間推定

5−1.591＜μ_{B1}＜5＋1.591　　3.409＜μ_{B1}＜6.591

● B_2水準の母平均の区間推定

9−1.591＜μ_{B2}＜9＋1.591　　7.409＜μ_{B2}＜10.591

● B_3水準の母平均の区間推定

11.5−1.591＜μ_{B3}＜11.5＋1.591　　9.909＜μ_{B3}＜13.091

● B_4水準の母平均の区間推定

6.5−1.591＜μ_{B4}＜6.5＋1.591　　4.909＜μ_{B4}＜8.091

手順4　最適条件での母平均の推定

母平均μの点推定を以下の式から求めます。$\widehat{\mu_{A1B3}}$＝A_1水準の平均値＋B_3水準の平均値−総平均値＝9.5＋11.5−8＝13.0

母平均μの区間推定（信頼度95％）

母平均の区間推定$\widehat{\mu_{A1B3}}$を以下の式から求めます。

$\widehat{\mu_{A1B3}}-t(\phi_e, 0.05)\sqrt{V_e/n_e} \leqq \mu_{A1B3} \leqq \widehat{\mu_{A1B3}}+t(\phi_e, 0.05)\sqrt{V_e/n_e}$

n_eは「有効反復係数」で、次の式で求められます。

有効反復係数$n_e=ab/(1+\phi_A+\phi_B)$（田口の公式）

（a：Aの水準数、b：Bの水準数）

（有効反復係数と田口の公式についてはP.142参照）

n_eを計算すると、$n_e=ab/(1+\phi_A+\phi_B)=2\times4/(1+1+3)=1.6$

$t(\phi e, 0.05)\sqrt{V_e/n_e}$を計算します。$t(\phi e, 0.05)\sqrt{V_e/n_e}=t(3, 0.05)\times\sqrt{1/1.6}\fallingdotseq3.182\times0.791\fallingdotseq2.52$　　$13.0-2.52\leqq\mu_{A1B3}\leqq13.0+2.52$

$10.48\leqq\mu_{A1B3}\leqq15.52$

6）二元配置実験（繰り返しあり）

2つの因子A、Bについて、それぞれa個の水準、b個の水準を選び、全部で$a\times b$個の組み合わせの実験をランダムに繰り返して行います。

水準数$a=4$、$b=3$、繰り返し数2回とすると、A_iの第i水準、B_jの第j水準を組み合わせた水準の下で行った実験のデータx_{ijk}は、次の構造式で観測されると考え、データ表は以下のとおりとなります。

データ＝総平均　＋処理の効果　＋交互作用　＋誤差

x_{ijk}　＝　μ　　$+\alpha_i+\beta_j$　　$+(\alpha\beta)_{ij}$　　$+\varepsilon_{ijk}$

繰り返しのある場合は、要因A、Bの主効果だけでなく、交互作用も調べることができます。

データ表

	B_1	B_2	B_3
A_1	x_{111}	x_{211}	x_{311}
	x_{112}	x_{212}	x_{312}
A_2	x_{121}	x_{221}	x_{321}
	x_{122}	x_{222}	x_{322}
A_3	x_{131}	x_{231}	x_{331}
	x_{132}	x_{232}	x_{332}
A_4	x_{141}	x_{241}	x_{341}
	x_{142}	x_{242}	x_{342}

要因	平方和S	自由度ϕ	平均平方V	分散比F_0
因子A	$S_A=\Sigma\dfrac{(A_i\text{のデータの合計})^2}{A_i\text{のデータ数}}-CT$	ϕ_A＝水準数−1	$V_A=\dfrac{S_A}{\phi_A}$	$F_0=\dfrac{V_A}{V_e}$
因子B	$S_B=\Sigma\dfrac{(B_j\text{のデータの合計})^2}{B_j\text{のデータ数}}-CT$	ϕ_B＝水準数−1	$V_B=\dfrac{S_B}{\phi_B}$	$F_0=\dfrac{V_B}{V_e}$
交互作用$A\times B$	$S_{A\times B}=S_{AB}-S_A-S_B$ $S_{AB}=\Sigma\dfrac{(AB\text{二元表の各数値})^2}{\text{繰り返し数}}-CT$	$\phi_{A\times B}=\phi_A\times\phi_B$	$V_{A\times B}=\dfrac{S_{A\times B}}{\phi_{A\times B}}$	$F_0=\dfrac{V_{A\times B}}{V_e}$
誤差e	$S_e=S_T-S_A-S_B-S_{A\times B}$	$\phi_e=\phi_T-\phi_A-\phi_B-\phi_{A\times B}$	$V_e=\dfrac{S_e}{\phi_e}$	
合計	$S_T=\Sigma$（データの二乗）$-CT$	ϕ_T＝総データ数−1		

変動の分解

総変動（総平方和）	因子Aによる変動	A因子の級間平方和S_A
	因子Bによる変動	B因子の級間平方和S_B
	交互作用$A\times B$による変動	交互作用$A\times B$の平方和$S_{A\times B}$
	級内変動	誤差平方和Se

データ表（因子Aを４水準、因子Bを３水準の実験）

	B_1	B_2	B_3
A_1	9	10	9
	9	10	8
A_2	9	11	7
	8	10	6
A_3	8	11	9
	6	12	10
A_4	5	9	8
	7	9	9

①分散分析

手順１ データの二乗表、二元表、二元表の二乗表の作成

データの二乗表

	B_1	B_2	B_3	総計
A_1	81	100	81	262
	81	100	64	245
A_2	81	121	49	251
	64	100	36	200
A_3	64	121	81	266
	36	144	100	280
A_4	25	81	64	170
	49	81	81	211
合計	481	848	556	1885

データの二元表

因子A ＼ 因子B	B_1	B_2	B_3	総計
A_1	18	20	17	55
A_2	17	21	13	51
A_3	14	23	19	56
A_4	12	18	17	47
合計	61	82	66	209

※データ表の同じ枠内の数値を足し合わせた表（A_1B_1は９＋９＝18）

データの二元表の二乗表

因子A ＼ 因子B	B_1	B_2	B_3	総計
A_1	324	400	289	1013
A_2	289	441	169	899
A_3	196	529	361	1086
A_4	144	324	289	757
合計	953	1694	1108	3755

手順２ 修正項（CT）の計算

CT＝（データの合計）^{2}／データ数＝209^2／24≒**1820.04**

手順３ 分散分析表の作成（平方和、自由度、平均平方、分散比の計算）

まず、各平方和を計算します。

$S_T=\Sigma$（データの二乗）$-CT=1885-1820.04=$**64.96**

$$S_A=\Sigma\frac{(A_i\text{のデータの合計})^2}{A_i\text{のデータ数}}-CT=\frac{55^2+51^2+56^2+47^2}{6\,(=3\times2)}-1820.04$$

$=1828.5-1820.04=$**8.46**

$$S_B=\Sigma\frac{(B_j\text{のデータの合計})^2}{B_j\text{のデータ数}}-CT=\frac{61^2+82^2+66^2}{8\,(=4\times2)}-1820.04$$

$\fallingdotseq1850.13-1820.04=$**30.09**

$$S_{AB}=\Sigma\frac{(AB\text{二元表の各数値})^2}{\text{繰り返し数}}-CT$$

$$=\frac{18^2+17^2+14^2+12^2+20^2+21^2+23^2+18^2+17^2+13^2+19^2+17^2}{2}-1820.04$$

$=\dfrac{3755}{2}-1820.04=1877.5-1820.04=$**57.46**

$S_{A\times B}=S_{AB}-S_A-S_B=57.46-8.46-30.09=$**18.91**

$S_e=S_T-S_A-S_B-S_{A\times B}=64.96-8.46-30.09-18.91=$**7.50**

次に、各自由度を計算します。

ϕ_T＝総データ数－１＝24－１＝**23**

ϕ_A＝水準数－１＝４－１＝**3**

ϕ_B＝水準数－１＝３－１＝**2**

$\phi_{A\times B}=\phi_A\times\phi_B=3\times2=$**6**

$\phi_e=\phi_T-\phi_A-\phi_B-\phi_{A\times B}=23-3-2-6=$**12**

続いて、平均平方と分散比を計算します。

$V_A = S_A / \phi_A = 8.46 / 3 = 2.82$

$V_B = S_B / \phi_B = 30.09 / 2 \fallingdotseq 15.05$

$V_{A\times B} = S_{A\times B} / \phi_{A\times B} = 18.91 / 6 \fallingdotseq 3.15$

$V_e = S_e / \phi_e = 7.50 / 12 \fallingdotseq 0.63$

これにより、分散比は、

A ： $F_0 = V_A / V_e = 2.82 / 0.63 \fallingdotseq 4.48$

B ： $F_0 = V_B / V_e = 15.05 / 0.63 = 23.89$

$A\times B$ ： $F_0 = V_{A\times B} / V_e = 3.15 / 0.63 = 5.00$

よって、分散分析表は次の表のとおりとなります。

要因	平方和S	自由度ϕ	平均平方V	分散比F_0
因子A	8.46	3	2.82	4.48
因子B	30.09	2	15.05	23.89
交互作用$A\times B$	18.91	6	3.15	5.00
誤差e	7.50	12	0.63	
合計	64.96	23		

手順4 分散分析結果の判定

手順３で得た分散比A：$F_0=4.48$、B：$F_0=23.89$、$A\times B$：$F_0=5.00$とF表の$F(\phi_A, \phi_e; \alpha) = F(3, 12; 0.05) = 3.49$、$F(\phi_B, \phi_e; \alpha) = F(2, 12; 0.05) = 3.89$、$F(\phi_{A\times B}, \phi_e; \alpha) = F(6, 12; 0.05) = 3.00$を比較すると、

A ： $F_0 = 4.48 > F(3, 12; 0.05) = 3.49$

B ： $F_0 = 23.89 > F(2, 12; 0.05) = 3.89$

$A\times B$ ： $F_0 = 5.00 > F(6, 12; 0.05) = 3.00$

となるので、有意な差があると判定できます。

判定には３つのパターンがあります。

①交互作用$A\times B$が有意でなく、因子Aまたは因子Bの片方が有意

②交互作用$A\times B$が有意でなく、因子Aおよび因子Bの両方が有意

③交互作用$A\times B$が有意(因子Aおよび因子Bが有意であるかどうかにかかわらず)

※本例は③に該当する。

①②の場合、プーリング(交互作用$A\times B$の平方和と自由度を誤差項のそれに加え込むこと)を行い、推定精度を上げることができます。

プーリングとは？

プーリングとは、効果のない項を誤差と見なして、それらの平方和と自由度を誤差項の平方和と自由度に足し込み、新たな誤差分散を求めることです。

これにより、誤差の自由度が増え、誤差分散の推定精度が上がります。

⇒ 具体的には、交互作用$A\times B$の平方和($S_{A\times B}$)と自由度($\phi_{A\times B}$)を誤差項Eの平方和(S_e)と自由度(ϕ_e)にそれぞれ加え込むことです。これにより分散分析表を作り直します。

要因	平方和S	自由度ϕ	平均平方V	分散比F_0
因子A	S_A	ϕ_A	V_A	$F_0=\frac{V_A}{Ve}$
因子B	S_B	ϕ_B	V_B	$F_0=\frac{V_B}{Ve}$
交互作用$A\times B$	$S_{A\times B}$	$\phi_{A\times B}$	$V_{A\times B}$	$F_0=\frac{V_{A\times B}}{Ve}$
誤差e	Se	ϕe	Ve	
合計	S_T	ϕ_T		

有意でなく、F_0値も小さい場合

プーリング

要因	平方和S	自由度ϕ	平均平方V	分散比F_0
因子A	S_A	ϕ_A	V_A	$F_0=\frac{V_A}{Ve'}$
因子B	S_B	ϕ_B	V_B	$F_0=\frac{V_B}{Ve'}$
誤差e'	$Se'=S_{A\times B}+Se$	$\phi e'=\phi_{A\times B}+\phi e$	Ve'	$F_0=\frac{V_{A\times B}}{Ve'}$
合計	S_T	ϕ_T		

交互作用$A\times B$を誤差項にプーリング

②推定

手順1 最適な組み合わせ条件の選定

特性値が大きい方がよいので、二元表において最大値となるA_3B_2を選定します。

手順2 最適条件での母平均の推定

母平均μの点推定

$\hat{\mu}(A_3B_2) = A_3B_2$の平均値$= (11+12)/2 = 23/2 = 11.5$

母平均μの区間推定(信頼度95%)

$$\widehat{\mu_{AiBj}} - t(\phi_e, 0.05)\sqrt{\frac{Ve}{n}} < \mu_{AiBj} < \widehat{\mu_{AiBj}} + t(\phi_e, 0.05)\sqrt{\frac{Ve}{n}}$$ (n：繰り返し数)

なお、交互作用$A\times B$が有意でなく、プーリングを行う場合(交互作用$A\times B$を無視する場合)、以下の式により区間推定を行う。

$$\widehat{\mu_{AiBj}} - t(\phi_e', 0.05)\sqrt{Ve'/n_e} < \mu_{AiBj} < \widehat{\mu_{AiBj}} + t(\phi_e', 0.05)\sqrt{Ve'/n_e}$$

※因子と水準の組み合わせで、特性値が最も大きくなる条件を「最適条件」という

有効反復係数 $n_e = \dfrac{abn}{1 + \phi_A + \phi_B + \phi_{A\times B}}$ （田口の公式）

（a ：Aの水準数、 b ：Bの水準数、 n ：繰り返し数）

で表されます。

$t(\phi_e, 0.05)\sqrt{V_e/n}$ を計算します。

$t(\phi_e, 0.05)\sqrt{\dfrac{V_e}{n}} = t(12, 0.05)\sqrt{\dfrac{0.63}{2}} \fallingdotseq 2.179\times 0.561 \fallingdotseq$ 1.22

$11.5 - 1.22 < \mu_{A3B2} < 11.5 + 1.22$

$10.28 < \mu_{A3B2} < 12.72$

有効反復係数 n_eと田口の公式、伊奈の公式

有効反復係数 n_eとは、点推定量が何個分のデータから計算されたものと等価であるかを示すものです。有効反復係数 n_eを計算する主な公式として、田口の公式と伊奈の公式があります。

田口の公式：$n_e = \dfrac{\text{全データ数}}{1 + (\text{推定に用いる要因の自由度の和})}$

(二元配置分析では) $= \dfrac{abn}{1 + \phi_A + \phi_B + \phi_{A\times B}}$

（a ：Aの水準数、 b ：Bの水準数、 n ：繰り返し数）

伊奈の公式：$\dfrac{1}{n_e} = $ 点推定の式に用いられている係数の和 $= \dfrac{1}{a} + \dfrac{1}{b} - \dfrac{1}{ab}$

練習問題

赤シートで正解を隠して設問に答えてください（解説はP.145から）。

【問１】 ３つの工場を選定し、ある品質特性における繰り返し３回の一元配置実験を行ったところ、以下の統計量が得られた。この統計量をもとに下表の分散分析表を作成することになった。$\boxed{\quad}$に入る数値を答えよ。ただし、(９)についてはもっとも適切なものを下の選択肢からひとつ選べ。

個々のデータの合計値 $\Sigma\Sigma x_{ij} = 250$

個々のデータの二乗の合計値 $\Sigma\Sigma x_{ij}^2 = 7500$

各水準の合計値の二乗の総和 $= \Sigma T_i^2 = 20000$

分散分析表

要因	平方和S	自由度ϕ	平均平方V	分散比F_0
因子A	(１)	(３)	(６)	(８)
誤差E	100	(４)	(７)	−
計	(２)	(５)	−	−

以上より、因子Aは (９) といえる。

〈(１)～(８)の選択肢〉

ア. 556　**イ**. 456　**ウ**. 228　**エ**. 16.7　**オ**. 13.65

カ. 8　**キ**. 6　**ク**. 2

〈(９)の選択肢〉

ア. 有意な差がある　イ. 有意な差がない

正解　(１) イ　(２) ア　(３) ク　(４) キ　(５) カ
　(６) ウ　(７) エ　(８) オ　(９) ア

【問２】 次の①②に答えよ。

①データの構造式において、Ａ群とＢ群を正しく組み合わせよ。

Ａ群：

ア. 一元配置法

イ. 二元配置法（繰り返しなし）

ウ. 二元配置法（繰り返しあり）

Ｂ群：

エ. $x_{ij} = \mu + \alpha_i + \varepsilon_{ij}$

オ. $x_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$

カ. $x_{ijk} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$

※ΣΣの計算方法を紹介する。$\sum_{i=1}^{n}\sum_{j=1}^{m} x_{ij} = x_{11} + x_{12} + \cdots + x_{1m} + x_{21} + x_{22} + \cdots + x_{2m} + \cdots + x_{n1} + x_{n2} + \cdots + x_{nm}$　まずは $i=1$ として、j を１からmまで動かしてx_{1j}を足し合わせる。すなわち、Σx_{1j}を求める。続いて、この操作を $i = 2, 3, \cdots, m$まで同様に繰り返して、すべてを足し合わせる

【問２】要因配置実験

①アーエ、イーカ、ウーオ

A群		B群
ア．一元配置法	—	エ．$x_{ij}=\mu+\alpha_i+\varepsilon_{ij}$
イ．二元配置法（繰り返しなし）	—	カ．$x_{ijk}=\mu+\alpha_i+\beta_j+\varepsilon_{ij}$
ウ．二元配置法（繰り返しあり）	—	オ．$x_{ijk}=\mu+\alpha_i+\beta_j+(\alpha\beta)_{ij}+\varepsilon_{ijk}$

②アーク、イーオ、ウーカ、エーキ

A群		B群
ア．因子	—	ク．実験に取り上げられた要因
イ．水準	—	オ．因子を量的、質的に変動させる場合に、特に代表値として選んだ値
ウ．効果	—	カ．応答の平均に対する因子の影響
エ．要因	—	キ．ある結果を引き起こす可能性のあるもの

【問３】二元配置実験

（１）ア　（２）ウ　（３）カ

二元配置法における因子による効果には、実験に取り上げた因子の水準を変更することによる（１）ア．主効果と、二つ以上の因子のある水準が組み合わさったときに相乗的または相殺的に特性値が変動する（２）ウ．交互作用効果の２種類がある。交互作用効果を検定するには、各水準の組み合わせの実験を最低（３）カ．２回繰り返す必要がある。

Ⅰ 手法編

第７章

抜取検査

合格のポイント

- ➡ 抜取検査の考え方・手順、用語の意味、ＯＣ曲線（検査特性曲線）の見方の理解
- ➡ 計数規準型抜取検査および計量規準型一回抜取検査の基本的な考え方や検査表の見方の理解

手順4 **試料の大きさ n と合格判定個数 c を求める**

- p_0、p_1の値から、計数規準型抜取検査表を用いて n と c を求める。
- n がロットの大きさを超える場合は全数検査を行う。

手順5 **試料の採取**

- 検査ロットの中から、手順4で求めた大きさ n の試料をできるだけロットを代表するようにしてとる。

手順6 **試料の調査**

- 手順1で定めた品質基準に従って試料を調べ、試料中の不良品の数を調べる。

手順7 **判定**

- 試料中の不良品の数が合格判定個数 c 以下であればそのロットを合格とし、c を超せばそのロットを不合格とする。

手順8 **ロットの処置**

- 合格または不合格と判定されたロットは、あらかじめ決めた約束に従って処置する。どのような場合でも、不合格となったロットをそのままで再提出してはならない。

〈例1〉P_0=1.0%、P_1=5.0%の場合の n、c を求める。

付表(計数規準型1回抜き取り検査表)において、P_0=1.0%を含む行と、P_1=5.0%を含む列の交わる欄の数値を読み取る。欄内には「120　3」とあり、左側が n、右側が c の値となる。よって、n=120、c=3となる。

P_0(%) \ P_1(%)	0.71~0.90	0.91~1.12	1.13~1.40	1.41~1.80	1.81~2.24	2.25~2.80	2.81~3.55	3.56~4.50	4.51~5.60	5.61~7.10	7.11~9.00	9.01~11.2	11.3~14.0
0.090~0.112	*	400 1	↓	←	↓	→	60 0	50 0	←	↓	↓	←	↓
0.113~0.140	*	↓	300 1	↓	←	↓	→	↑	40 0	←	↓	↓	←
0.141~0.180	*	500 2	↓	250 1	↓	←	↓	→	↑	30 0	←	↓	↓
0.181~0.224	*	*	400 2	↓	200 1	↓	←	↓	→	↑	25 0	←	↓
0.225~0.280	*	*	500 3	300 2	↓	150 1	↓	←	↓	→	↑	20 0	←
0.281~0.355	*	*	*	400 3	250 2	↓	120 1	↓	←	↓	→	↑	15 0
0.356~0.450	*	*	*	500 4	300 3	200 2	↓	100 1	↓	←	↓	→	↑
0.451~0.560	*	*	*	*	400 4	250 3	150 2	↓	80 1	↓	←	↓	→
0.561~0.710	*	*	*	*	500 6	300 4	200 3	120 2	↓	60 1	↓	←	↓
0.711~0.900	*	*	*	*	*	400 6	250 4	150 3	100 2	↓	50 1	↓	←
0.901~1.12		*	*	*	*	*	300 6	200 4	120 3	80 2	↓	40 1	↓
1.13 ~1.40			*	*	*	*	500 10	250 6	150 4	100 3	60 2	↓	30 1
1.41 ~1.80				*	*	*	*	400 10	200 6	120 4	80 3	50 2	↓
1.81 ~2.24					*	*	*	*	300 10	150 6	100 4	60 3	40 2

〈例2〉P_0=0.15%、P_1=2.5%の場合の n、c を求める。

付表(計数規準型1回抜き取り検査表)において、P_0=0.15%を含む行と、P_1=2.5%を含む列の交わる欄の数値を読み取る。欄内には「←」とあるので、数値の入った欄まで左側に進む。「250　1」とあり、左側が n、右側が c の値となることから、n=250、c=1となる。

P_0(%) \ P_1(%)	0.71~0.90	0.91~1.12	1.13~1.40	1.41~1.80	1.81~2.24	2.25~2.80	2.81~3.55	3.56~4.50	4.51~5.60	5.6~7.1
0.090~0.112	*	400 1	↓	←	↓	→	60 0	50 0	←	↓
0.113~0.140	*	↓	300 1	↓	←	↓	→	↑	40 0	←
0.141~0.180	*	500 2	↓	250 1	↓	←	↓	→	↑	30

〈例3〉P_0=0.3%、P_1=1.0%の場合の n、c を求める。

付表(計数規準型1回抜き取り検査表)において、P_0=0.3%を含む行と、P_1=1.0%を含む列の交わる欄の数値を読み取る。欄内には「＊」とあるので、付表(抜取検査設計補助表)を参照し、P_1／P_0=1.0／0.3≒3.33を含む3.5~2.8の行にある数値となることから、c=6、n=164／P_0+527／P_1=164／0.3+527／1.0=1073.67→1074(抜き取り検査の値は正の整数となるので、小数点以下を切り上げる)。

P_0(%) \ P_1(%)	0.71~0.90	0.91~1.12	1.13~1.40
0.090~0.112	*	400 1	↓
0.113~0.140	*	↓	300 1
0.141~0.180	*	500 2	↓
0.181~0.224	*	*	400 2
0.225~0.280	*	*	500 3
0.281~0.355	*	*	*

4 計量規準型一回抜取検査(標準偏差既知)

標準偏差既知でロットの平均値を保証する場合および標準偏差既知でロットの不良率を保証する場合の計量規準型一回抜取検査(JIS Z 9003-1979)とは、ロットの品質をロットの平均値または不良率で表した場合に生産者および消費者の要求する検査特性を持つように設計した抜取検査であって、**一回に抜き取った試料の特性値の平均値に対し既知の標準偏差を使って計算した合格判定値を比較することによって、ロットの合格・不合格を判定する**ものです。

この検査の適用に当たっては、

(1)検査単位の品質は、計量値で表し得ること
(2)製品がロットとして処理できること
(3)ロットの特性値の標準偏差がわかっていることが必要であること
(4)不良率による場合は、特性値が正規分布をしているものとして取り扱われ

2 管理図の種類

利用目的・取り扱うデータによって、**管理図**を使い分けます。以下に主な管理図を紹介します。データに偏りが大きいときは、平均値ではなく、中央値を用いる場合があります。

図表8.3 管理図の種類

特性	名称	概要 (評価値→評価されるもの)	中心線(CL)、上方管理限界線(UCL)、下方管理限界線(LCL)	
計量値／連続量(正規分布を仮定する)	$\overline{X}-R$ 管理図	$\overline{X}$管理図とR管理図を組み合わせたもの。サンプル数が一定で、9以下の場合に用いる。 $\overline{X}$：データ群の平均を時系列で打点したもので、工程水準の評価に用いる R：各データ群の最大値と最小値の差(範囲)を時系列で打点したもので、変動の評価に用いる。	【$\overline{X}$管理図】 $CL=\overline{\overline{x}}$ $UCL=\overline{\overline{x}}+A_2\overline{R}$ $LCL=\overline{\overline{x}}-A_2\overline{R}$ ※母平均μ_0と母標準偏差σ_0が与えられている場合は、 $CL=\mu_0$ $UCL=\mu_0+A\sigma_0$ $LCL=\mu_0-A\sigma_0$	【R管理図】 $CL=\overline{R}$ $UCL=D_4\overline{R}$ $LCL=D_3\overline{R}$
	$\overline{X}-s$ 管理図	$\overline{X}$管理図とs管理図を組み合わせたもの。サンプル数が一定で、10以上の場合に用いる。 $\overline{X}$：データ群の平均を時系列で打点したもので、工程水準の評価に用いる s：各データ群の標準偏差(バラツキ具合)を打点したもので、変動の評価に用いる。	【$\overline{X}$管理図】 $CL=\overline{\overline{x}}$ $UCL=\overline{\overline{x}}+A_3\overline{s}$ $LCL=\overline{\overline{x}}-A_3\overline{s}$	【s管理図】 $CL=\overline{s}$ $UCL=B_4\overline{s}$ $LCL=B_3\overline{s}$
	$Me-R$ 管理図	Me管理図とR管理図を組み合わせたもの。データの平均に偏りがある場合に用いる。 Me：データ群の中央値を時系列で打点したもので、工程水準の評価に用いる R：各データ群の最大値と最小値の差(範囲)を時系列で打点したもので、変動の評価に用いる。	【Me管理図】 $CL=\overline{Me}$ $UCL=\overline{Me}+A_4\overline{R}$ $LCL=\overline{Me}-A_4\overline{R}$ ($Me=\widetilde{x}$の平均値) ($A_4=m_3A_2$ P.165参照)	【R管理図】 $CL=\overline{R}$ $UCL=D_4\overline{R}$ $LCL=D_3\overline{R}$
計数値／離散値(二項分布を仮定する)	np 管理図 (number：数、propotion：割合)	サンプルサイズ(n)が一定の場合に用いる。全体の合計(n)×不適合品率(p)＝不適合品数→不適合品数を打点したもので、不適合品数の評価に用いる。	$CL=n\overline{p}$ $UCL=n\overline{p}+3\sqrt{n\overline{p}(1-\overline{p})}$ $LCL=n\overline{p}-3\sqrt{n\overline{p}(1-\overline{p})}$	
	p 管理図 (propotion：割合)	サンプルサイズが変動し、サンプルサイズを一定にできないときに用いる。 不適合品数／サンプル数の不適合品率を打点したもので、不良品率(不適合品率)の評価に用いる。	$CL=\overline{p}$ $UCL=\overline{p}+3\sqrt{\overline{p}(1-\overline{p})/n}$ $LCL=\overline{p}-3\sqrt{\overline{p}(1-\overline{p})/n}$	
計数値／離散値(ポアソン分布を仮定する)	c 管理図 (count：数える)	サンプルサイズが一定の場合に用いる。 不適合数(c)を打点したもので、不適合品数の評価に用いる。	$CL=\overline{c}$ $UCL=\overline{c}+3\sqrt{\overline{c}}$ $LCL=\overline{c}-3\sqrt{\overline{c}}$	
	u 管理図 (unit：単位)	サンプルサイズが変動し、一定でない場合に用いる。 不適合数／サンプル数の単位あたりの不適合率(u)を打点したもので、単位当たりの不適合品数の評価に用いる。	$CL=\overline{u}$ $UCL=\overline{u}+3\sqrt{\overline{u}/n}$ $LCL=\overline{u}-3\sqrt{\overline{u}/n}$	

※A_1、A_2、A_3、B_3、B_4、D_3、D_4はサンプルサイズnによって定まる(各係数表参照)

3 管理図の作成手順

1）$\overline{X}-R$ 管理図 ……平均値と範囲の管理図

サンプル数が一定で、９以下の場合に用いられます。

手順１ データの収集・データ表の作成

群の番号	測定値		計	平均値	範囲
	x_1	x_2	Σx	$\overline{x}$	R
1	5	6	11	5.5	1
2	4	5	9	4.5	1

手順２ 管理値の計算

①平均値

群ごとに平均値を計算する。

群１の平均値 $\overline{x}_1$＝（５＋６）／２＝**5.5**

群２の平均値 $\overline{x}_2$＝（４＋５）／２＝**4.5**

②範囲

群ごとに範囲 R を計算する。

R＝x の最大値－x の最小値

群１の範囲 R_1＝６－５＝**1**

群２の範囲 R_2＝５－４＝**1**

図表8．4 $\overline{X}$-R 管理図用係数表

サンプルサイズ	$\overline{X}$ 管理図		R 管理図			
n	A	A_2	D_3	D_4	d_2	d_3
2	2.121	1.88	—	3.27	1.128	0.853
3	1.732	1.02	—	2.57	1.693	0.888
4	1.500	0.73	—	2.28	2.059	0.880
5	1.342	0.58	—	2.11	2.326	0.864
6	1.225	0.48	—	2.00	2.534	0.848
7	1.134	0.42	0.08	1.92	2.704	0.833
8	1.061	0.37	0.14	1.86	2.847	0.820
9	1.000	0.34	0.18	1.82	2.970	0.808
10	0.949	0.31	0.22	1.78	3.078	0.797

手順３ 管理限界線の計算

● $\overline{X}$管理図

中心線 $CL=\overline{\overline{x}}=(5.5+4.5)/2=$**5.0**　※ $\overline{\overline{x}}$ は $\overline{x}$ の平均値

上方管理限界線 $UCL=\overline{\overline{x}}+A_2\overline{R}=5.0+1.88\times1=$**6.88**

下方管理限界線 $LCL=\overline{\overline{x}}-A_2\overline{R}=5.0-1.88\times1=$**3.12**

（A_2は係数表から求める）

● R管理図

上方管理限界線 $UCL=D_4\overline{R}=3.27\times1=$**3.27**

下方管理限界線 $LCL=D_3\overline{R}$　※ $n\leqq6$ の場合は考慮しない

（D_4、D_3は付表から求める。サンプルサイズ $n=2$）

2）$\overline{X}-s$ 管理図 ……平均値と標準偏差（バラツキ）の管理図

サンプル数が一定で、10以上の場合に用いられます。

手順１ データの収集・データ表の作成

群番号	測定値										計	平均値	標準偏差
	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	Σx	$\overline{x}$	s
1	5	6	2	3	4	1	5	2	4	5	37	3.7	1.64
2	4	5	1	4	5	6	2	2	3	1	33	3.3	1.77

手順２ 管理値の計算

①平均値

群ごとに平均値を計算する。

群１の平均値 $\overline{x}_1$＝（５＋６＋２＋３＋４＋１＋５＋２＋４＋５）／10＝**3.7**

群２の平均値 $\overline{x}_2$＝（４＋５＋１＋４＋５＋６＋２＋２＋３＋１）／10＝**3.3**

②標準偏差

群ごとに標準偏差 s を計算する。

$$s=\sqrt{V}=\sqrt{\frac{S}{n-1}}=\sqrt{\frac{\Sigma(x_i-\overline{x})^2}{n-1}}=\sqrt{\frac{\Sigma x_i^2-(\Sigma x_i)^2/n}{n-1}}$$

標準偏差 $s_1=\sqrt{\dfrac{161-37^2/10}{10-1}}\fallingdotseq$**1.636**

標準偏差 $s_2=\sqrt{\dfrac{137-33^2/10}{10-1}}\fallingdotseq$**1.767**　➡ $\overline{s}=\dfrac{1.636+1.767}{2}\fallingdotseq$**1.70**

図表8.5 $\overline{X}$-s 管理図用係数表

サンプルサイズ	$\overline{X}$管理図	s管理図	
n	A_3	B_4	B_3
2	2.659	3.267	—
3	1.954	2.568	—
4	1.628	2.266	—
5	1.427	2.089	—
6	1.287	1.970	0.030
7	1.182	1.882	0.118
8	1.099	1.815	0.185
9	1.032	1.761	0.239
10	0.975	1.716	0.284

手順3 管理限界線の計算

● $\overline{X}$管理図

中心線 $CL=\overline{\overline{x}}=(3.7+3.3)/2=3.5$

上方管理限界線 $UCL=\overline{\overline{x}}+A_3\bar{s}=3.5+0.975\times1.70\fallingdotseq5.16$

下方管理限界線 $LCL=\overline{\overline{x}}-A_3\bar{s}=3.5-0.975\times1.70\fallingdotseq1.84$

（A_3は係数表から求める）

● s管理図

中心線 $CL=\bar{s}=1.70$

上方管理限界線 $UCL=B_4\times\bar{s}=1.716\times1.70\fallingdotseq2.92$

下方管理限界線 $LCL=B_3\times\bar{s}=0.284\times1.70\fallingdotseq0.48$

（B_4、B_3は付表から求める）

3）MeーR管理図 ……中央値と範囲の管理図

データの平均に偏りがあり、**平均値を用いることが好ましくない場合**に用いられます。Me はMedian（メディアン、中央値）の略で、$\tilde{x}$ と表記されることもあります。

手順1 データの収集・データ表の作成

群番号	測定値										計	中央値	範囲
	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	Σx	$\tilde{x}$	R
1	5	6	2	3	4	1	5	2	4	5	37	4	5
2	4	5	1	4	5	6	2	2	3	1	33	3.5	5

手順2 管理値の計算

①中央値

群ごとに中央値を計算する。

群１の中央値 $\tilde{x}_1=4$

群２の中央値 $\tilde{x}_2=(3+4)/2=3.5$

②範囲

群ごとに範囲 R を計算する。

$R=x$ の最大値 $-\ x$ の最小値

群１の範囲 $R_1=6-1=5$

群２の範囲 $R_2=6-1=5$

図表8.6 Me管理図係数表

サンプルサイズn	Me管理図
	$m_3A_2(=A_4)$
2	1.880
3	1.187
4	0.796
5	0.691
6	0.549
7	0.509
8	0.432
9	0.412
10	0.363

手順3 管理限界線の計算

● Me管理図

中心線 $CL=\overline{\tilde{x}}=(4+3.5)/2=3.75$　※ $\overline{\tilde{x}}$ は $\tilde{x}$ の平均値

上方管理限界線 $UCL=\overline{\tilde{x}}+m_3A_2\overline{R}=3.75+0.363\times5=5.565$

下方管理限界線 $LCL=\overline{\tilde{x}}-m_3A_2\overline{R}=3.75-0.363\times5=1.935$

※ $m_3A_2(=A_4)$ は群の大きさ n によって定まる値で、**Me管理図係数表**から求める

● R管理図

中心線 $CL=\overline{R}=5$

上方管理限界線 $UCL=D_4\overline{R}=1.78\times5=8.90$

下方管理限界線 $LCL=D_3\overline{R}=0.22\times5=1.10$

（D_4、D_3は係数表から求める。P.162参照）

4) np管理図 ……不適合品数の管理図

不適合品数について用います。群の大きさが一定である必要があります。

手順1 **データの収集・データ表の作成**

○日目	不適合品数np
1	1
2	2
3	0
4	1
5	3

n＝20個とする。

手順2 **管理値の計算**

● **平均不適合品数$n\bar{p}$**

$n\bar{p}=\Sigma(np)_i/k=(1+2+0+1+3)/5=1.4$

$(np)_i$：各群の不適合品数、k：群の数

なお、$\bar{p}$（平均不適合品率）は、$\bar{p}=\Sigma(np)_i/\Sigma n_i=\dfrac{7}{100}$

手順3 **管理限界線の計算**

中心線$CL=n\bar{p}=1.4$

上方管理限界線$UCL=n\bar{p}+3\sqrt{n\bar{p}(1-\bar{p})}$

$=1.4+3\sqrt{1.4(1-7/100)}\fallingdotseq 4.82$

下方管理限界線$LCL=n\bar{p}-3\sqrt{n\bar{p}(1-\bar{p})}$

$=1.4-3\sqrt{1.4(1-7/100)}\fallingdotseq -2.02$

マイナスになる場合は考慮しない。

5) p管理図 ……不適合品率の管理図

不適合品率について用います。群の大きさが一定でない場合に用います。

手順1 **データの収集・データ表の作成**

○日目	群の大きさ	不適合品数
1	50	1
2	50	2
3	30	0
4	30	1
5	40	3

不適合品率（p）は、不適合品数（np）／群の大きさ（n）により計算する。

手順2 **管理値の計算**

● **平均不適合品率$\bar{p}$**

$\bar{p}=\Sigma(np)_i/\Sigma n_i=\dfrac{1+2+0+1+3}{50+50+30+30+40}=0.035$

$(np)_i$：各群の不適合品数、n_i：群の大きさ

手順3 **管理限界線の計算**

中心線$CL=\bar{p}=0.035$

n_i＝50のとき、

上方管理限界線$UCL=\bar{p}+3\sqrt{\bar{p}(1-\bar{p})/n_i}$

$=0.035+3\sqrt{0.035(1-0.035)/50}\fallingdotseq 0.11$

下方管理限界線$LCL=\bar{p}-3\sqrt{\bar{p}(1-\bar{p})/n_i}$

$=0.035-3\sqrt{0.035(1-0.035)/50}\fallingdotseq -0.04$

マイナスになる場合は考慮しない。

n_i＝30のとき、

上方管理限界線$UCL=\bar{p}+3\sqrt{\bar{p}(1-\bar{p})/n_i}$

$=0.035+3\sqrt{0.035(1-0.035)/30}\fallingdotseq 0.14$

下方管理限界線$LCL=\bar{p}-3\sqrt{\bar{p}(1-\bar{p})/n_i}$

$=0.035-3\sqrt{0.035(1-0.035)/30}\fallingdotseq -0.07$

マイナスになる場合は考慮しない。

n_i＝40のとき、

上方管理限界線$UCL=\bar{p}+3\sqrt{\bar{p}(1-\bar{p})/n_i}$

$=0.035+3\sqrt{0.035(1-0.035)/40}\fallingdotseq 0.12$

下方管理限界線$LCL=\bar{p}-3\sqrt{\bar{p}(1-\bar{p})/n_i}$

$=0.035-3\sqrt{0.035(1-0.035)/40}\fallingdotseq -0.05$

マイナスになる場合は考慮しない。

6) c 管理図 ……各群の不適合数の管理図

群の大きさが一定のサンプルを採取して、**各群における欠点数（不適合数）を調査**します。

手順1 データの収集・データ表の作成

群No.	群の大きさ	欠点数	群No.	群の大きさ	欠点数
1	50	1	11	50	1
2	50	2	12	50	2
3	50	0	13	50	0
4	50	1	14	50	1
5	50	3	15	50	3
6	50	4	16	50	0
7	50	1	17	50	2
8	50	0	18	50	3
9	50	2	19	50	1
10	50	3	20	50	4

欠点数(不適合数)(c)は、欠点数の総和(Σc)／群の大きさ(k)により計算する。

手順2 管理値の計算

● **平均欠点数(不適合数) $\bar{c}$**

$\bar{c}=\Sigma c／k=(1+2+0+1+3+4+1+0+2+3+1+2+0+1+3+0+2+3+1+4)／20=34／20=$**1.7**

$\bar{c}$ ：工程平均欠点数、Σc ：欠点数の総和、k ：群の数

手順3 管理限界線の計算

中心線 $CL=\bar{c}=1.7$

上方管理限界線 $UCL=\bar{c}+3\sqrt{\bar{c}}$

$=1.7+3\sqrt{1.7}\fallingdotseq$ **5.61**

下方管理限界線 $LCL=\bar{c}-3\sqrt{\bar{c}}$

$=1.7-3\sqrt{1.7}\fallingdotseq$ **−2.21**

マイナスになる場合は考慮しない。

7) u 管理図 ……単位あたりの不適合率の管理図

群の大きさの異なるサンプルを採取して、**サンプルサイズ(面積、長さ等)とサンプル中の欠点数(不適合数)を調査**します。

手順1 データの収集・データ表の作成

群No.	群の大きさ	欠点数	群No.	群の大きさ	欠点数
1	50	1	11	30	1
2	50	2	12	30	2
3	50	0	13	30	0
4	50	1	14	30	1
5	50	3	15	30	3
6	40	4	16	20	0
7	40	1	17	20	2
8	40	0	18	20	3
9	40	2	19	20	1
10	40	3	20	20	4

手順2 管理値の計算

● **平均欠点率(不適合率) $\bar{u}$**

$\bar{u}=\Sigma c／\Sigma n=(1+2+0+1+3+4+1+0+2+3+1+2+0+1+3+0+2+3+1+4)／(50\times5+40\times5+30\times5+20\times5)=34／700\fallingdotseq$**0.049**

$\bar{u}$ ：工程平均欠点率、Σc ：欠点数の総和、Σn ：サンプルサイズの総和

手順3 管理限界線の計算

中心線 $CL=\bar{u}=$**0.049**

$n_i=50$のとき、

上方管理限界線 $UCL=\bar{u}+3\sqrt{\bar{u}/n_i}$

$=0.049+3\sqrt{0.049/50}\fallingdotseq$**0.14**

下方管理限界線 $LCL=\bar{u}-3\sqrt{\bar{u}/n_i}$

$=0.049-3\sqrt{0.049/50}\fallingdotseq$**−0.04**

マイナスになる場合は考慮しない。

$n_i=40$のとき、

上方管理限界線 $UCL=\bar{u}+3\sqrt{\bar{u}/n_i}$

$=0.049+3\sqrt{0.049/40}\fallingdotseq$**0.15**

下方管理限界線 $LCL=\bar{u}-3\sqrt{\bar{u}/n_i}$

$=0.049-3\sqrt{0.049/40}\fallingdotseq$**−0.06**

マイナスになる場合は考慮しない。

$n_i=30$のとき、

上方管理限界線 $UCL=\bar{u}+3\sqrt{\bar{u}/n_i}$

$=0.049+3\sqrt{0.049/30}\fallingdotseq$**0.17**

1 工程能力指数

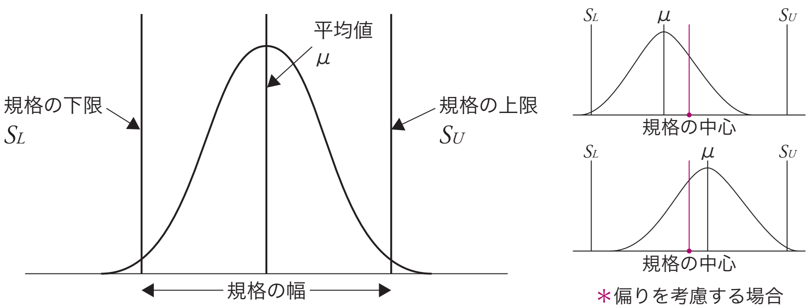
工程能力とは、**定められた規格の限度内で製品を生産できる能力**のことです。(JIS Z8101:2015において)統計的管理状態(異常原因が取り除かれた安定状態)にあることが実証されたプロセスについての、特性の成果に関する統計的推定値であり、プロセスが特性に関する要求事項を実現する能力を記述したものと定義されています。

この評価を行う指標を**工程能力指数**といい、C_p(C_{pk})(Process Capability Index)と表されます。

C_pとC_{pk}との使い分けについて、C_pは品質特性値の母平均が規格の中心にあることがわかっている場合に使用されます。一方、C_{pk}はそうでない(偏りを考慮する)場合に使用されます。

母集団分布が正規分布$N(\mu,\ \sigma_2)$に従うと仮定した場合、工程能力指数の算定式は以下のとおりです。

図表9.1 平均値、規格の上限・下限、規格の幅



1）両側規格の場合

上限規格値S_Uと下限規格値S_Lの両方が決まっている場合は、以下の式で工程能力を算定します。

$$C_p = \frac{S_U - S_L}{6\sigma} \qquad (\sigma：標準偏差)$$

2）片側規格の場合

上限規格値S_Uまたは下限規格値S_Lの**いずれか片方のみが決まっている場合**は、以下の式で工程能力を算定します。

①上限規格値S_Uのみが決まっている場合

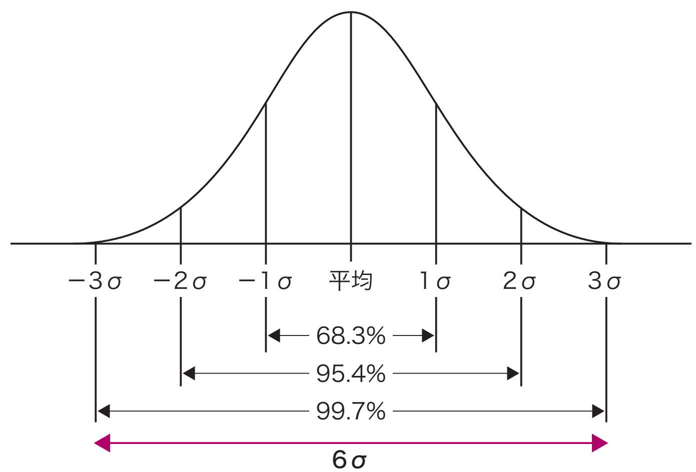
$$C_p(上限) = \frac{S_U - \mu}{3\sigma} \qquad (\sigma：標準偏差、\mu：平均値)$$

②下限規格値S_Lのみが決まっている場合

$$C_p(下限) = \frac{\mu - S_L}{3\sigma} \qquad (\sigma：標準偏差、\mu：平均値)$$

なお、両側規格の場合でも**平均値が規格の中心(規格の幅／2)から大きくズレている(偏りがある)場合**があります。このような場合はC_{pk}と表記し、上記①と②の計算を行い、①と②から得られた値を比較して、小さい方を採用します。

図表9.2 $N(\mu,\ \sigma^2)$において$\mu \pm a\sigma$に入る確率



なお、標準偏差は、範囲Rを用いて、$\hat{\sigma} = \overline{R} / d_2$で示されます。
(d_2：係数(管理限界線を計算するための係数表から求める))

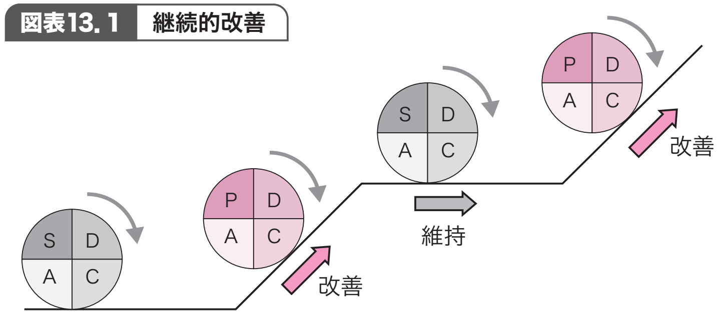
1 維持と管理、継続的改善

維持とは、**製品・サービスの適切な状態をそのまま保ち続けること**であり、維持のために管理を適切に行う必要があります。具体的には**5M**（原材料(Material)、設備(Machine)、方法(Method)、作業者(Man)、計測(Measurement)）を用いて、**SDCAサイクル**を回して管理を行います。

一方、製品・サービスに対する、顧客のニーズ・期待は常に進化することから、**継続的改善**を行う必要があります。

継続的改善とは、製品・サービスの品質向上のために繰り返し行われる活動であり、具体的には**SDCAサイクル**と**PDCAサイクル**を組み合わせて行っていきます。

図表13.1　継続的改善

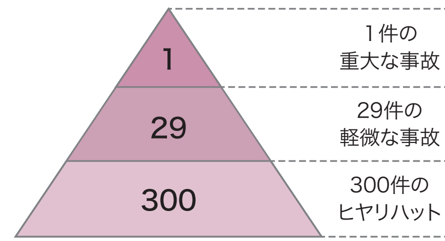


S：Standardize
D：Do
C：Check
A：Act

P：Plan
D：Do
C：Check
A：Act

SDCA	Standardize（標準化）→Do（実行）→Check（評価）→Act（措置）の頭文字を取ったもの。日常管理を行ううえでのサイクル。
PDCA	Plan（計画）→Do（実行）→Check（評価）→Act（措置）の頭文字を取ったもの。改善活動を行ううえでのサイクル。

その他の重要キーワードは以下の通りです。

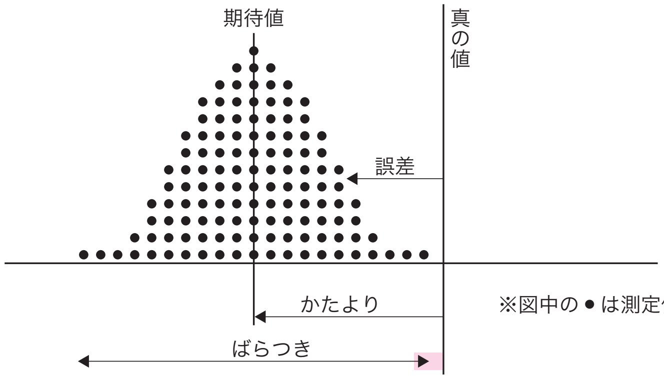
キーワード	概　要
ハインリッヒの法則	ハインリッヒ（米国損害保険会社勤務）が5,000件以上の労働災害の調査結果から得た経験則である。重大な災害・事故が発生したケースにおいて、その背景には29件の小さな事故やケガが発生しており、300件のヒヤリハットがあったというものである。
見える化	プロセスの状況を関係者全員が理解できるように可視化すること。

2 自主保全活動

自主保全活動とは、作業者一人一人が自分の担当する設備は自分で守ることを目的として、**7つのステップ**で設備の保全活動を行うものです。

ステップ	名　称	主な内容
第1ステップ	初期清掃	設備のゴミ・ホコリ・汚れの清掃を行い、設備の不具合発見と復元を行う。
第2ステップ	発生源・困難箇所の対策	ゴミ・ホコリ・汚れの発生源、清掃・給油・増締め・点検等が困難な箇所の改善により、清掃・点検時間を短縮する。
第3ステップ	自主保全仮基準の作成	清掃・給油・増締めを短時間で確実に行う行動基準類を作成する。

図表14.4 誤差・かたより・ばらつき



※図中の●は測定値を示す。

練習問題 赤シートで正解を隠して設問に答えてください(解説はP.222から)。

【問１】 品質保証に関する次の文章において、［　　］内に入る最も適切なものを下の選択肢からひとつ選べ。

新製品開発において、量産段階での不具合発生による手戻り・ロスをできるだけ防ぐため、［（1）］での問題予測・是正・改善が重要である。

製品の設計・開発を進めるうえで、設計者・開発者の知見だけでなく、より多くの関係者(営業、製造、購買、品質保証部門等)を集め、知見や［（2）］を結集し、さまざまな角度から検討することを［（3）］という。［（3）］には、製品企画への移行の可否を決める［（4）］審査、製品設計への移行の可否を決める［（5）］審査、設計品質達成状況を確認するための［（6）］審査がある。

〈選択肢〉

ア．ＤＲ	イ．ＦＭＥＡ	ウ．ＦＴＡ	エ．源流段階
オ．検査段階	カ．力量	キ．情報	ク．製品企画
ケ．商品企画	コ．試作設計		

正解　（1）エ　（2）カ　（3）ア　（4）ケ　（5）ク　（6）コ

【問２】 品質保証に関する次の文章において、［　　］内に入る最も適切なものを下の選択肢からひとつ選べ。

ＱＣ工程図は、［（1）］とも呼ばれ、［（2）］方向に工程名、作業内容、作業標準書、管理項目等の情報が記載され、［（3）］方向に製品の加工、調整、検査等のプロセスを記載したものである。製品の設計段階で定めた仕様が［（4）］段階で実現できることが狙いの一つであり、［（4）］現場の管理方法を示す指示文書としての役割も果たす。管理方法についてはできるだけ［（5）］的に記載する。

〈選択肢〉

ア．ＱＦＤ	イ．ＱＣ工程表	ウ．ＱＡネットワーク
エ．縦	オ．横	カ．企画
キ．生産	ク．抽象	ケ．具体

正解　（1）イ　（2）オ　（3）エ　（4）キ　（5）ケ

【問３】 品質保証に関する次の文章において、［　　］内に入る最も適切なものを下の選択肢からひとつ選べ。

①品質管理を行うために、ヒストグラムを用いられることが多い。ヒストグラムを作成するにあたり、各々の寸法等の［（1）］と、［（2）］を示す標準偏差を算出する。また、上限と下限の［（3）］を記入することで、工程能力等を視覚的に把握することができる。

〈選択肢〉

ア．中央値	イ．平方和	ウ．平均値	エ．ばらつき
オ．度数	カ．管理限界線	キ．ＯＣ曲線	

正解　①（1）ウ　（2）エ　（3）カ

②多数のデータを要因によっていくつかのグループに分けることを［（1）］という。これは［（2）］において８つめの道具と呼ばれることも多い。これにより、例えば、是正措置や改善措置の前後のデータを比較し、各措置の［（3）］性を判断することができる。

生産した製品に何か不具合が生じた場合に、追跡調査ができる状態にしてお

【問７】 実験計画法

実験計画法に関する次の文章において、□内に入るもっとも適切なものを下のそれぞれの選択肢からひとつ選び、その記号を解答欄にマークせよ。ただし、各選択肢を複数回用いることはない。なお、解答にあたって必要であれば巻末の付表を用いよ。

製品Xの品質特性の調査のため、品質に影響すると考えられる因子Aについて調べることになった。因子Aを３水準設定し、繰り返し２回で計６回の実験をランダムに行ったところ、下表の結果を得た。

表

因子A	A_1	A_2	A_3
データ	5.5	4.9	7.6
	5.7	4.5	6.9
合計	11.2	9.4	14.5
合計の２乗	125.44	88.36	210.25
個々のデータの２乗和	62.74	44.26	105.37

①得られたデータを用いて、分散分析表を作成すると下表のとおりとなる。

要因	平方和S	自由度ϕ	平均平方V	分散比F_0
因子A	(1)	(3)	(5)	(7)
誤差e	(2)	(4)	(6)	
計	7.04	5		

【(１)～(７)の選択肢】

ア．0.12　イ．0.35　ウ．3.35　エ．6.69　オ．29.09
カ．1　キ．2　ク．3　ケ．4

②分散分析表およびF分布表より、品質は (8) 。また、データの数値が大きいほど望ましいとすると、もっとも望ましい数値は (9) のときである。

【(８)～(９)の選択肢】

ア．有意でない　イ．有意である　ウ．A_1　エ．A_2　オ．A_3

【問８】 単回帰分析

単回帰分析に関する次の文章において、□内に入るもっとも適切なものを下のそれぞれの選択肢からひとつ選び、その記号を解答欄にマークせよ。ただし、各選択肢を複数回用いてもよい。

ある工程の要因xと品質特性yに関する対のあるデータを10組観測した結果、次の統計量の値を得た。

xの平均値：$\bar{x}=6.4$、yの平均値：$\bar{y}=8.4$、xの偏差平方和：$S_{xx}=66.21$、yの偏差平方和：$S_{yy}=21.43$、xとyの偏差積和：$S_{xy}=34.56$

①要因xと品質特性yの相関係数rは、 (1) となる。

②xを説明変数とし、yを目的変数として回帰式($y=a+bx$)を推定すると、回帰係数$b=$ (2) となる。

③②の推定された回帰式について分散分析表を作成すると下表のとおりとなる。

要因	平方和S	自由度ϕ	平均平方V	分散比F_0
回帰	(3)	(5)	(7)	(9)
残差	(4)	(6)	(8)	
計	21.43	9		

【(１)～(９)の選択肢】

ア．0.42　イ．0.52　ウ．0.92　エ．3.39　オ．18.04
カ．42.95　キ．1　ク．2　ケ．7　コ．8

F表より、$F(\phi_R,\ \phi_e\,;0.05)=$ (10) なので、回帰は (11) 。

【(10)～(11)の選択肢】

ア．有意でない　イ．有意である　ウ．4.74　エ．5.32