

判定ルールと不確かさ 理解のための統計セミナー

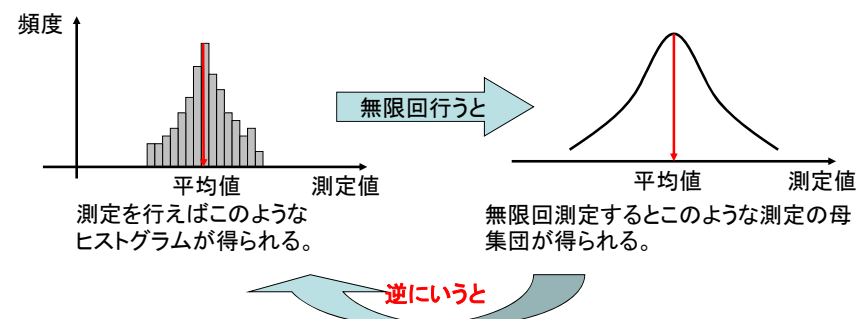
産業技術総合研究所 計量標準総合センター
田中秀幸

1: 統計の基本的事項について

統計学の初歩

1-1: 統計から見た測定とは？

測定と統計について

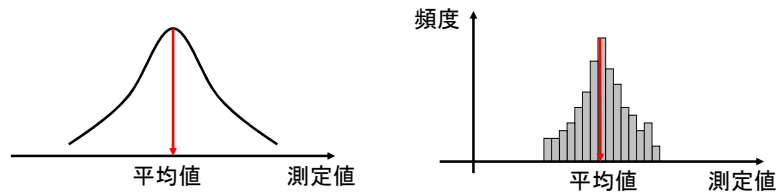


測定をしようと決めたときに、すでに母集団の形は決まっている。
測定を行えばその母集団から値をひとつ取り出すということになる。

測定とは、その測定を無限回繰返したとき得られる母集団からのサンプリングである。

母集団と標本について(1)

私たちは測定することによってなにを知りたいのだろう？



本当ならこの母集団の性質を知りたい。しかし、母集団の性質も理想的な量であるので知ることができない。

そこで、無限回測定する事は無理なので有限回の測定によって母集団の性質を推定する。

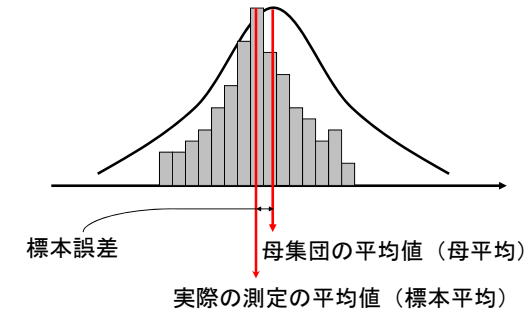
サンプルの性質を知りたいわけではない

一番知りたいもの

母集団の性質

母集団と標本について(2)

一般的に代表値としてよく用いられている平均値を考えてみる。



このように、本当に知りたい値である母平均は、サンプルから求められた標本平均とずれがある。統計の重要な目的は、この母平均の推定値とその推定された値がどの程度正しいのかということ知ることである。

数式・数字の表記方法(統計)

標本を表すもの: アルファベット(確率変数という)

→ 確率変数...ある確率に従って値が変わる変数

例: サイコロの目, 測定値, ...

母集団の性質を表すもの(母数): ギリシャ文字(定数)

ギリシャ文字の上に「^」(ハット)が付くもの: 推定値

i 番目の測定結果: x_i

標本平均: $\bar{x}$

母平均: μ

標本平均を母平均の推定値とする:

$$\hat{\mu} = \bar{x}$$

連続分布と離散分布

離散分布...飛び飛びの値しか取らない要素からなる分布

例: サイコロの目の値, 個数, 人数....

連続分布...離散分布とは違い, ある区間内に値が存在しているとしても, その区間内で無限個のデータが含まれている要素からなる分布

例: 測定値

離散分布は概念や統計処理において理解しやすいが, 連続分布は無限の概念を考える必要があり, 理解しにくい。

ただし, 測定値は通常連続分布であるので, 連続分布における統計処理について知る必要がある。

1-2: 離散分布の期待値について

期待値とは

- 期待値・・・理想的にはこの値になるという値
つまりサイコロを1回振ったときこれくらいの値
が出ることが期待できる,という値のこと。
サイコロは1-6の面を一つずつ持っている。
また, その6つの面がでる確率はそれぞれ1/6。
よって期待値は,

$$1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6}$$

$$= \frac{1}{6} + \frac{2}{6} + \frac{3}{6} + \frac{4}{6} + \frac{5}{6} + \frac{6}{6} = \frac{21}{6} = 3.5$$

期待値とは

- 一般的に書くと

$$E(x) = x_1 P(x_1) + x_2 P(x_2) + \cdots + x_n P(x_n)$$

$$E(x) = \sum_{i=1}^n x_i P(x_i)$$

期待値は, $E(x)$ と書く。

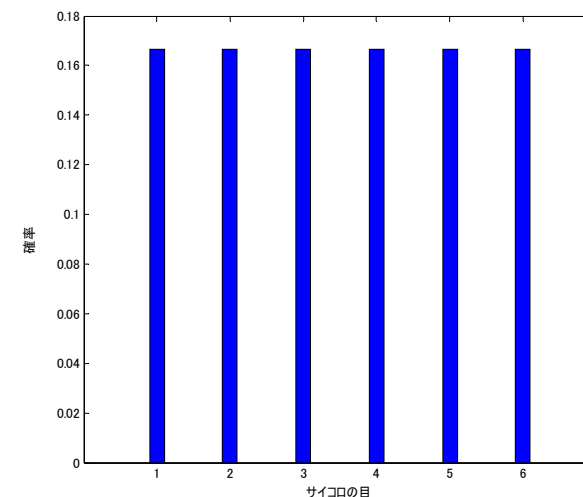
また, $P(x_i)$ のことを確率関数という。確率変数 x の値が x_i となる
確率を表す関数。

$$P(x) = \frac{1}{6} \quad \{x \mid x=1, 2, 3, 4, 5, 6\}$$

例: サイコロの確率関数

$$= 0 \quad \{x \mid x=1, 2, 3, 4, 5, 6 \text{ 以外の整数}\}$$

サイコロの確率関数のグラフ



期待値例：ルーレット

ルーレットは1-36までの穴と、0と00の穴がある。
また、賭けた番号に玉が入ると36倍の配当が貰える。
この時、私は\$100をある番号に賭けました。
貰える配当の期待値はいくらか？

解答：ルーレット

- 穴の個数は1-36の穴と0と00の穴なので、38個。
- また、どの穴も入る確率は同じ。
よって、当たる確率は、 $1/38$ となる。

賭けた穴に玉が入った場合には、36倍の配当で\$3600貰える。その他の場合は、\$0となる。

解答：ルーレット

- 期待値は、

$$E(x) = 3600 \times \frac{1}{38} + 0 \times \frac{1}{38} + \cdots + 0 \times \frac{1}{38}$$

$$E(x) = 94.74$$

よって、\$100賭けたときの期待値は\$94.74となる。
つまり、この差額が親の取り分。

期待値の性質

測定値 x の母平均を μ とすると、 $\mu = E(x)$

c が定数のとき、 $E(c) = c$ $E(cx) = cE(x)$

x, y が確率変数のとき、 $E(x \pm y) = E(x) \pm E(y)$

x, y が確率変数で
互いに独立のとき、 $E(xy) = E(x)E(y)$

期待値の性質

$$c \text{ が定数のとき, } E(c) = c \quad E(cx) = cE(x)$$

$$E(c) = c \quad c \text{ は確率的に値が変わらないので、いつでも } c \text{ となる。}$$

$$E(cx) = cE(x)$$

2,4,6,8,10,12のサイコロを考えると、期待値は $(2+4+6+8+10+12)/6=7$ となる。これは、普通のサイコロの期待値3.5の2倍である。

期待値の性質

$$x, y \text{ が確率変数のとき, } E(x \pm y) = E(x) \pm E(y)$$

	1	2	3	4	5	6	和
1	2	3	4	5	6	7	27
2	3	4	5	6	7	8	33
3	4	5	6	7	8	9	39
4	5	6	7	8	9	10	45
5	6	7	8	9	10	11	51
6	7	8	9	10	11	12	57

全和:252
 $E(x+y)=$
 $252 \div 36=7$

$E(x)+E(y)=$
 $3.5+3.5=7$

二つのサイコロを同時に振って出た目を足した値について考える。二つのサイコロの出目の和は2と12の間になり、その中心の値(期待値)は7である。この値は、一つだけサイコロを振ったときの期待値3.5を2つ足し算したときと同じである。

期待値の性質

$$x, y \text{ が確率変数で互いに独立のとき, } E(xy) = E(x)E(y)$$

	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	4	6	8	10	12
3	3	6	9	12	15	18
4	4	8	12	16	20	24
5	5	10	15	20	25	30
6	6	12	18	24	30	36
和	21	42	63	84	105	126

全和:441
 $E(xy)=$
 $441 \div 36=12.25$

$E(x)E(y)=$
 $3.5 \times 3.5=12.25$

互いに独立: 確率変数 x と y が互いに影響し合うことがなく、全く別個に値が決定する、ということを意味している。

期待値の計算例

6面体のサイコロの目の値: x
 8面体のサイコロの目の値: y

$$E(x) = 1 \cdot \frac{1}{6} + \dots + 6 \cdot \frac{1}{6} = 3.5 \quad E(y) = 1 \cdot \frac{1}{8} + \dots + 8 \cdot \frac{1}{8} = 4.5$$

6面体のサイコロの出た目を3倍した値と8面体のサイコロの出た目を2倍した値の和の期待値は、

$$E(3x + 2y) = E(3x) + E(2y) = 3E(x) + 2E(y) \\ = 3 \times 3.5 + 2 \times 4.5 = 19.5$$

標本平均の期待値

標本平均: $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$ の期待値を考える。

$$E(\bar{x}) = E\left(\frac{\sum_{i=1}^n x_i}{n}\right) = \frac{1}{n} E(x_1 + \dots + x_n)$$

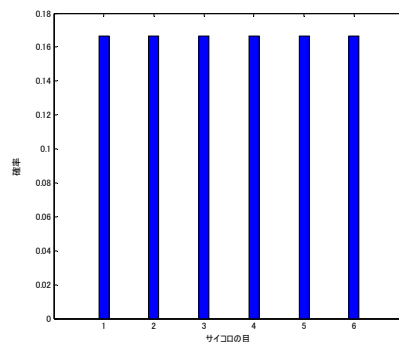
$$= \frac{1}{n} [E(x_1) + \dots + E(x_n)] = \frac{1}{n} \sum_{i=1}^n E(x_i) = \frac{1}{n} \sum_{i=1}^n \mu = \frac{1}{n} \cdot n\mu$$

$$E(\bar{x}) = \mu$$

標本平均は母平均の推定値として何らかのかたよりを持たない。このようにかたより無く母数の推定が行うことができる推定量のことを**不偏推定量**という。

1-3: 連続分布への拡張

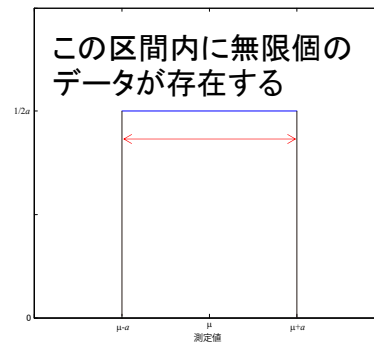
連続分布への拡張



離散分布

ある値を取るときの確率を求めることができる。

無限個からある要素一つがサンプリングされる確率は, $\frac{1}{\infty} = 0$



連続分布

ある値を取るときの確率を求めることが**できない**。

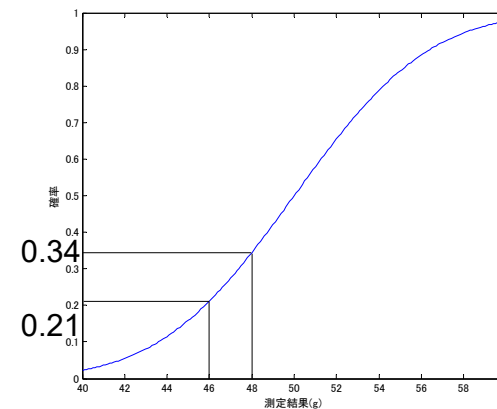
連続分布への拡張

では、連続分布のときは確率が計算できないのか？

→ある値になるときの確率が計算できないだけで、サンプリングされた値が、ある範囲に含まれる確率であれば計算できる。

測定結果が46gから48gの間に入る確率は、

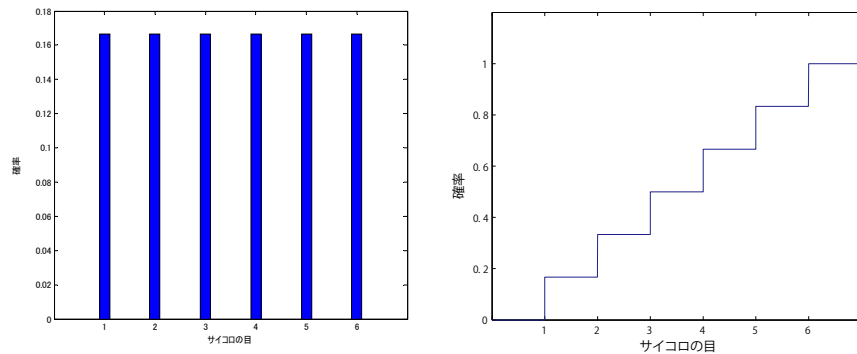
$$0.34 - 0.21 = 0.13$$



製品の質量の分布(**累積分布関数**)

左図で示したグラフを表す関数を**累積分布関数**または**分布関数**(Cumulative Distribution Function: **CDF**)といい、この分布関数は $F(x)$ で表す。

離散分布の累積分布関数



確率関数と累積分布関数の関係を考えると、確率関数では、1-6の目がすべて1/6の確率で出ており、累積分布関数では、目が1増える毎に1/6ずつ確率が増えることが分かる。つまり、**確率関数は累積分布関数の傾きを表している**ことになる。

確率密度関数

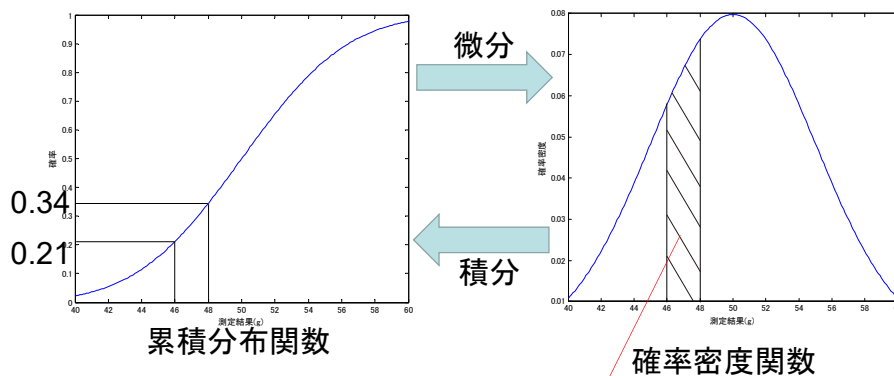
離散分布
確率関数は累積分布関数の傾きを表している



連続分布
累積分布関数の傾きを表した関数(つまり累積分布関数を微分したもの)が、連続分布における確率関数のようなものになる

この連続分布における確率関数に相当するものを
確率密度関数 (Probability Density Function: PDF)という

確率密度関数



測定結果が46gから48gの間に
入る確率は、 $0.34 - 0.21 = 0.13$

確率密度関数のグラフは直感的に母集団の分布を示しているの
で、一般的に確率分布を示すときには確率密度関数が使われる。

確率密度関数

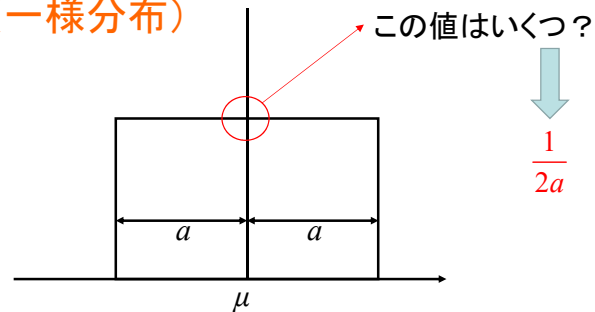
実は...

矩形分布(一様分布)	三角分布	U字分布	正規分布
最もよく使われる分布 限界値のときなどに 適用。	よく使われる分布。 中心が多く、端に いくほど少なくなる 分布に適用。	周期的に変化する 要因に対して適用。 「温度」「高周波の インピーダンスマッ チング」	校正証明書などで 不確かさが分かっ ている時に適用。
$\frac{a}{\sqrt{3}}$	$\frac{a}{\sqrt{6}}$	$\frac{a}{\sqrt{2}}$	$\frac{U}{k}$

ここで描かれているグラフはすべて確率密度関数！！

確率密度の値

矩形分布(一様分布)



確率密度と x 軸と任意の x の区間で挟まれる面積が確率となるので、 $-\infty < x < \infty$ 間で挟まれる面積は1となる必要がある。つまり、上記の長方形の面積が1となる。このとき底辺の長さが $2a$ であるので、高さは $1/(2a)$ となる。

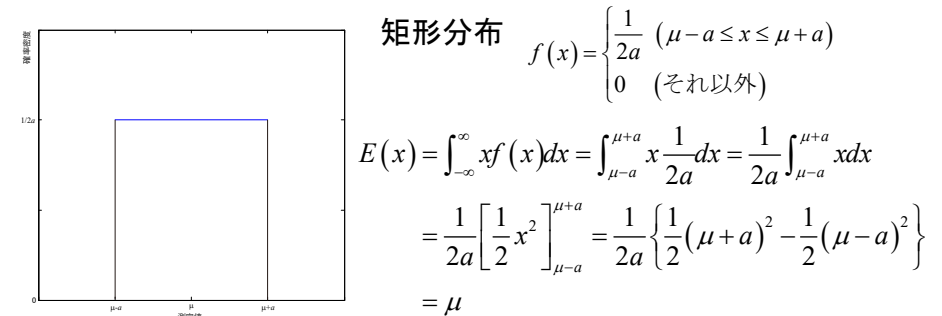
連続分布の期待値

離散分布

連続分布

$$E(x) = \sum_{i=1}^{\infty} x_i P(x_i) \quad E(x) = \int_{-\infty}^{\infty} x f(x) dx$$

測定値に(確率関数)(確率密度関数)を掛けて和を取る。



期待値の性質

測定値 x の母平均を μ とすると, $\mu = E(x)$

c が定数のとき, $E(c) = c$ $E(cx) = cE(x)$

x, y が確率変数のとき, $E(x \pm y) = E(x) \pm E(y)$

x, y が確率変数で互いに独立のとき, $E(xy) = E(x)E(y)$

1-4: ばらつきを表すパラメータ

母集団のばらつき

個々のデータが持つばらつきの大きさ: $\delta = x - \mu$ δ : 偏差

平均的なばらつきを知るために偏差の期待値を求める

$$E(\delta) = E(x) - \mu = \mu - \mu = 0$$

x は母平均より大きな値, 小さな値の両方取り得る可能性があるから

δ の大きさに着目する必要がある $\rightarrow$ δ を二乗する

$$E(\delta^2) = E[(x - \mu)^2]$$

この $E(\delta^2)$ のことを分散といい $V(x)$ で表す。 $V(x) = E[(x - \mu)^2]$

母分散について

$$V(x) = E[(x - \mu)^2]$$

$V(x)$ は x のばらつきを表すパラメータであるので、これを**母分散**といい σ^2 で表される。

なぜ σ に二乗が付くのかというと, 例えば, 測定結果が質量の[g]で表されていたとすると, この母分散の単位は偏差を二乗しているのので, $[g^2]$ となるからである。また, 二乗が付かない σ 単体で表される母数のことを**母標準偏差**という。

$$V(x) = E[(x - \mu)^2] = E(x^2 + \mu^2 - 2\mu x)$$

分散の別の表し方

$$\begin{aligned} &= E(x^2) + \mu^2 - 2\mu E(x) = E(x^2) - \mu^2 \\ &= E(x^2) - [E(x)]^2 \end{aligned}$$

母分散の計算法

離散分布

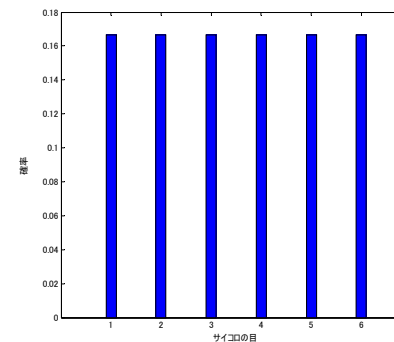
$$V(x) = E[(x - \mu)^2] = \sum_{i=1}^{\infty} (x_i - \mu)^2 P(x_i)$$

連続分布

$$V(x) = E[(x - \mu)^2] = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx$$

(測定値-母平均)²に(確率関数)(確率密度関数)を掛けて和を取る。

離散分布での母分散の計算



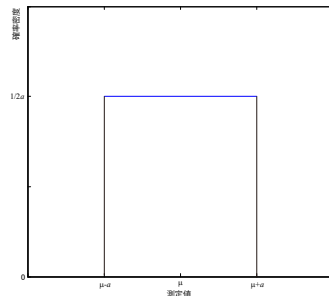
サイコロの目

サイコロの目の母分散は,

$$\begin{aligned} V(x) &= \sum_{i=1}^{\infty} (x_i - \mu)^2 P(x_i) \\ &= \frac{1}{6}(1-3.5)^2 + \frac{1}{6}(2-3.5)^2 + \\ &\quad \dots + \frac{1}{6}(6-3.5)^2 \\ &= 2.917 \end{aligned}$$

ちなみに, $\sqrt{2.917} = 1.708$

連続分布での母分散の計算



矩形分布

$$f(x) = \begin{cases} \frac{1}{2a} & (\mu - a \leq x \leq \mu + a) \\ 0 & (\text{それ以外}) \end{cases}$$

$$\begin{aligned} V(x) &= \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx \\ &= \int_{\mu-a}^{\mu+a} (x - \mu)^2 \frac{1}{2a} dx = \frac{1}{2a} \int_{\mu-a}^{\mu+a} (x - \mu)^2 dx \\ &= \frac{1}{2a} \left[\frac{1}{3} (x - \mu)^3 \right]_{\mu-a}^{\mu+a} = \frac{1}{2a} \left\{ \frac{1}{3} (a)^3 - \frac{1}{3} (-a)^3 \right\} \\ &= \frac{1}{2a} \cdot \frac{2a^3}{3} = \frac{a^2}{3} \end{aligned}$$

ちなみに、標準偏差で表せば $\frac{a}{\sqrt{3}}$

分散の性質

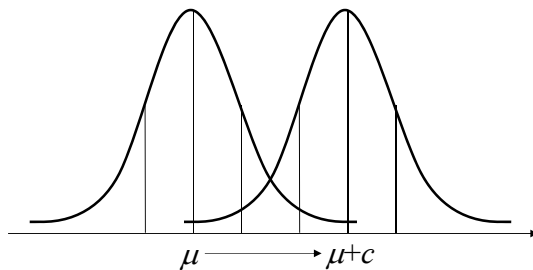
$$c \text{ が定数のとき, } V(x + c) = V(x) \quad V(cx) = c^2 V(x)$$

$$x, y \text{ が確率変数で互いに独立のとき, } V(x \pm y) = V(x) + V(y)$$

分散の性質

$$c \text{ が定数のとき, } V(x + c) = V(x)$$

x が c だけ平行移動したときのばらつきを考える。確率分布が平行移動しただけなら、分布の形自体は変わらないことからばらつきの大きさも変わらない。

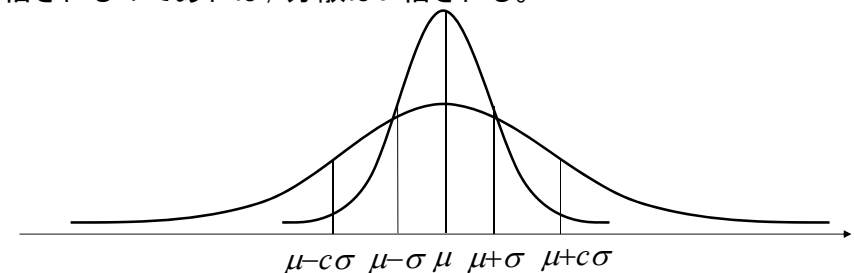


分散の性質

$$c \text{ が定数のとき, } V(cx) = c^2 V(x)$$

確率変数が c 倍されるのであるから、確率分布のばらつきは c 倍される。ただし、母分散の単位は測定対象量の単位の二乗となっているので、母集団のばらつきが c 倍されるのであれば、分散は c^2 倍される。

$$\begin{aligned} V(cx) &= E\left[\{c(x - \mu)\}^2\right] \\ &= E\left[c^2 (x - \mu)^2\right] \\ &= c^2 E\left[(x - \mu)^2\right] = c^2 V(x) \end{aligned}$$



分散の性質

x, y が確率変数で
互いに独立のとき,

$$V(x \pm y) = V(x) + V(y)$$

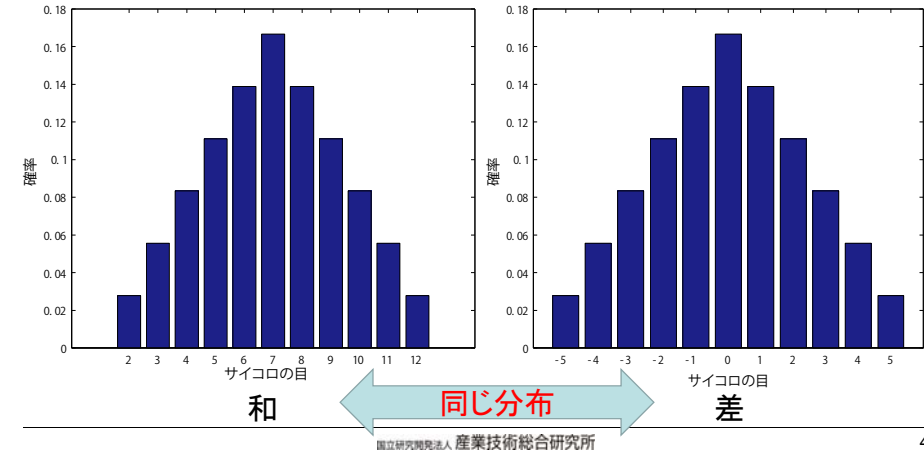
$$\begin{aligned} V(x \pm y) &= E\left[\{(x - \mu_x) \pm (y - \mu_y)\}^2\right] \\ &= E\left[(x - \mu_x)^2 + (y - \mu_y)^2 \pm 2(x - \mu_x)(y - \mu_y)\right] \\ &= E\left[(x - \mu_x)^2\right] + E\left[(y - \mu_y)^2\right] \pm 2E\left[(x - \mu_x)(y - \mu_y)\right] \\ E\left[(x - \mu_x)(y - \mu_y)\right] &\text{は確率変数 } x, y \text{ が独立であるとき,} \\ E\left[(x - \mu_x)(y - \mu_y)\right] &= 0 \end{aligned}$$

$$\begin{aligned} V(x \pm y) &= E\left[(x - \mu_x)^2\right] + E\left[(y - \mu_y)^2\right] \\ &= V(x) + V(y) \end{aligned}$$

分散の性質

x, y が確率変数で
互いに独立のとき,

$$V(x \pm y) = V(x) + V(y)$$



不確かさの合成法

$$V(x \pm y) = V(x) + V(y)$$

上記の式は、不確かさを合成する際に用いられている。

$$u_c(y) = \sqrt{u_1^2 + u_2^2 + \dots}$$

標準偏差(標準不確かさ)の2乗なので分散

分散の性質演算例

コップに水を入れて質量を測定し、更にコップだけの質量を測定し、その差を求め、水の質量を求めた。

$$Z = X - Y$$

コップ+水の質量を表す確率変数: x
コップの質量を表す確率変数: y
水の質量を表す確率変数: z

$$E(x) = \mu_x$$

$$V(x) = \sigma_x^2$$

$$E(y) = \mu_y$$

$$V(y) = \sigma_y^2$$

$$E(z) = E(x - y)$$

$$V(z) = V(x - y)$$

$$= E(x) - E(y)$$

$$= V(x) + V(y)$$

$$= \mu_x - \mu_y$$

$$= \sigma_x^2 + \sigma_y^2$$

1-5: 母分散と標本分散

母分散と標本分散

母分散は母数なので実際には知ることができない。よって、母分散を標本分散によって推定することを考える。

残差: $d_i = x_i - \bar{x}$ (母平均を標本平均で代用)

期待値を求める代わりに標本平均を求めて代用する。

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$$

s^2 が標本分散である。この標本分散によって、母分散を推定する。

標本分散と不偏推定量

標本分散は母分散の不偏推定量であるかを考える。

不偏推定量であればもちろん, $E(s^2) = V(x) = \sigma^2$ が成立する。

$$\begin{aligned} E(s^2) &= E\left[\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}\right] = \frac{1}{n} E\left[\sum_{i=1}^n (x_i^2 - 2x_i\bar{x} + \bar{x}^2)\right] \\ &= \frac{1}{n} E\left[\sum_{i=1}^n x_i^2 - 2\bar{x} \sum_{i=1}^n x_i + \bar{x}^2 \sum_{i=1}^n 1\right] = \frac{1}{n} E\left[\sum_{i=1}^n (x_i^2) - 2\bar{x} \cdot n\bar{x} + n\bar{x}^2\right] \\ &= \frac{1}{n} E\left[\sum_{i=1}^n (x_i^2) - n\bar{x}^2\right] = \frac{1}{n} \sum_{i=1}^n E(x_i^2) - E(\bar{x}^2) \\ E(x_i^2) &= E(x^2) \text{ と考えてよいので, } E(s^2) = \frac{1}{n} \sum_{i=1}^n E(x_i^2) - E(\bar{x}^2) \end{aligned}$$

標本分散と不偏推定量

ここで, $E(x^2)$ と $E(\bar{x}^2)$ について考える。

$$\begin{aligned} V(x) &= E[(x - \mu)^2] = E(x^2 + \mu^2 - 2\mu x) \\ &= E(x^2) + \mu^2 - 2\mu E(x) \\ &= E(x^2) - \mu^2 \end{aligned} \quad \longrightarrow \quad E(x^2) = \sigma^2 + \mu^2$$

$$\begin{aligned} V(\bar{x}) &= E[(\bar{x} - \mu)^2] = E(\bar{x}^2 + \mu^2 - 2\mu\bar{x}) \\ &= E(\bar{x}^2) + \mu^2 - 2\mu E(\bar{x}) \\ &= E(\bar{x}^2) - \mu^2 \end{aligned} \quad \longrightarrow \quad \begin{aligned} V(\bar{x}) &= E(\bar{x}^2) - \mu^2 \\ \frac{V(x)}{n} &= E(\bar{x}^2) - \mu^2 \\ E(\bar{x}^2) &= \frac{\sigma^2}{n} + \mu \end{aligned}$$

標本分散と不偏推定量

両式を代入

つまり、 $E(s^2) = V(x)$ は成立しない。

$$\begin{aligned} E(s^2) &= \frac{1}{n} \sum_{i=1}^n E(x_i^2) - E(\bar{x}^2) \\ &= \frac{1}{n} \sum_{i=1}^n (\sigma^2 + \mu^2) - \left(\frac{\sigma^2}{n} + \mu^2 \right) \\ &= \frac{1}{n} \cdot n(\sigma^2 + \mu^2) - \left(\frac{\sigma^2}{n} + \mu^2 \right) \\ &= \sigma^2 - \frac{\sigma^2}{n} \\ &= \frac{n-1}{n} \sigma^2 \end{aligned}$$

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$$

上式は母分散を $(n-1)/n$ 倍された値、すなわち母分散より少し小さな値を推定していることとなる。よって不偏推定量であるとは言えない。

標本分散と不偏推定量

標本分散を不偏推定量にするためには、先程の式を $n/(n-1)$ 倍すればよい

$$s^2(x) = \frac{n}{n-1} \cdot \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$$

上式で表した標本分散のことを**不偏分散**という。通常測定データの処理には不偏分散を用いるので、これ以降の本解説内では、 $s^2(x)$ はすべて上式の不偏分散を表し、 n で割った分散の意味では用いない。

標本分散以外のばらつきを表す指標

なぜ、標本分散以外のばらつきの指標が統計の教科書にはあまり載っていないのか？



標本分散以外のほとんどのばらつきを表す指標は不偏推定量ではない！

不偏推定量ではないばらつきを表す指標

・標本標準偏差

$$s(x) = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$$

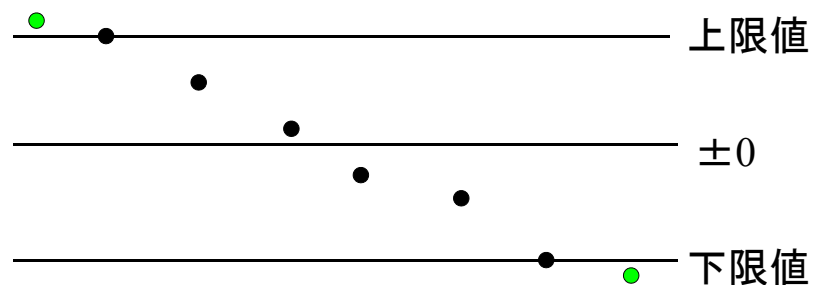
・平均偏差

$$M_{\text{dev}}(x) = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n}$$

2. 許容差と不確かさの関係

合否判定のリスクについて

従来の適合性表明の考え方

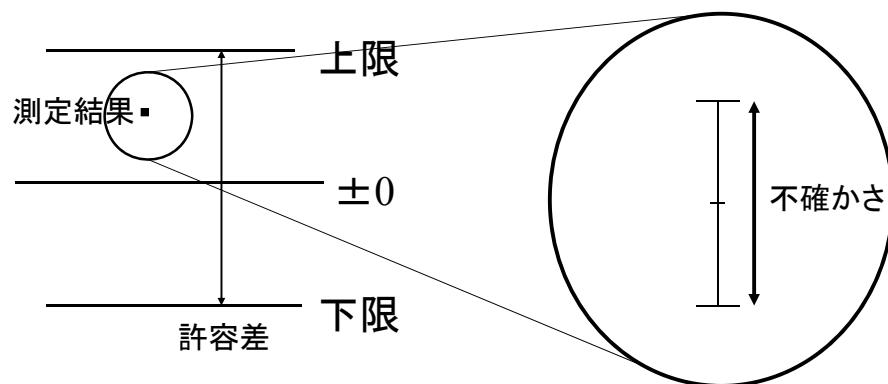


測定値が、上限値/下限値の間であれば、合格としていた。

許容差と不確かさの違い

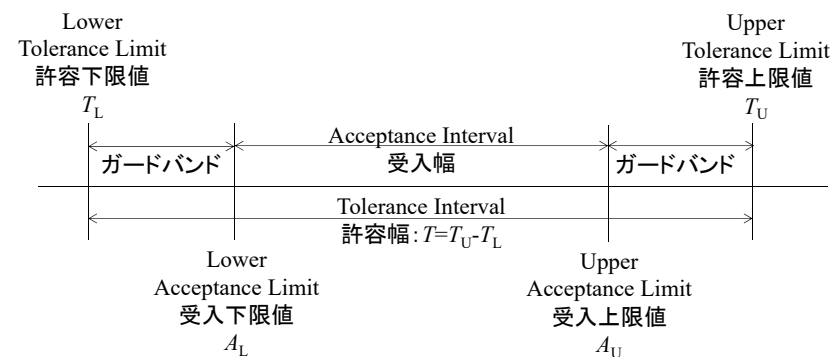
- 許容差 (製品規格, 最大許容誤差, 限界値)
合格判定を行うための基準。
- 不確かさ
測定結果は点で示されるものではなく
範囲を持っている, ということを表す。

許容差と不確かさの関係



許容差と不確かさは異なる概念！
でも密接な関係がある。

合否判定と不確かさに関する用語



注: 日本語訳は厳密なものではありません。

合否判定と不確かさに関する用語

- 許容限界(上限・下限)値・ T :この範囲の中に入っていれば合格, という判断基準。**この判断基準は生産システム, 測定システムの性質から決められるものではない!**これは, この範囲から外れると部品が組み込めない・不良品になる(製品スペックからの要件), この範囲から外れると健康被害が出る(これまでの知見・法律)など, 測定とは関係のないところから決められる。

合否判定と不確かさに関する用語

- 受入限界(上限・下限)値・ A :あるリスクを満たす前提で, 再設定される判断基準。**この判断基準は許容限界, 生産システム, 測定システムの性質から決められる!**
- ガードバンド: 許容限界値と受入限界値の差。場合によっては, 受入限界値が許容限界値の外側に設定されることもある。

合否判定と不確かさに関する用語

- 消費者リスク・・・本来は不合格のものを合格としてしまう確率。
- 生産者リスク・・・本来は合格のものを不合格としてしまう確率。

この二つのリスクは表裏一体の関係。つまり, 片方のリスクを小さくすると, もう片方のリスクが大きくなってしまう。

適合性評価についての文献

発行年	統計的手法	通則・品質システム	認定
1980		MIL-STD-45662A	
1998		ISO GPS規格(ISO14253-1, JIS B 0641-1)	
2001		ASME B89.7.3.1	
2003		ISO 10012	
2005	ASME B89.7.4.1		
2006		ANSI/NCSL Z540.3	
2009			ILAC G8
2010			APLAC TC004
2012	JCGM 106		
2017		OIML G 19	
2019			ILAC G8(改)

合否判定と不確かさの歴史

許容幅に含まれていれば合格、含まれていなければ不合格
(法定計量等幅広く用いられる。“Simple Acceptance”,
“Shared Risk”)

4:1ルールを導入(MIL規格)

許容幅に安全率・ガードバンドを経験的に設定(製品製造, 食品安全, 建築基準等幅広く用いられる)

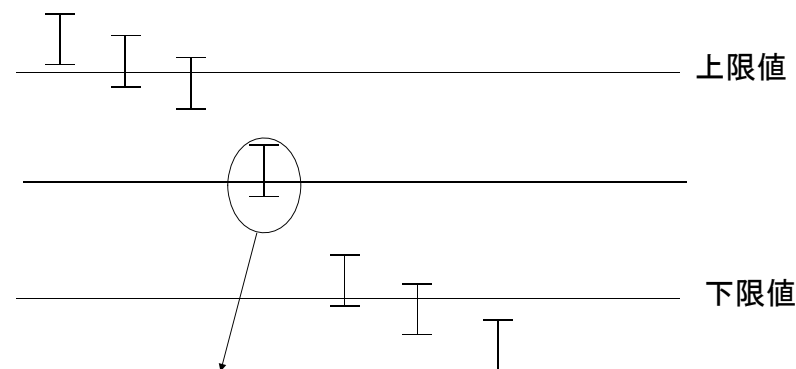
4:1ルールもしくは消費者リスク2%(Z540.3)

決めがなければガードバンドを拡張不確かさ分だけ設定(JIS B 0641-1)

望ましい消費者リスク, 生産者リスクを決定し, そのリスクを満たせるようにガードバンドを設定(ASME, JCGM106)

リスクとガードバンド

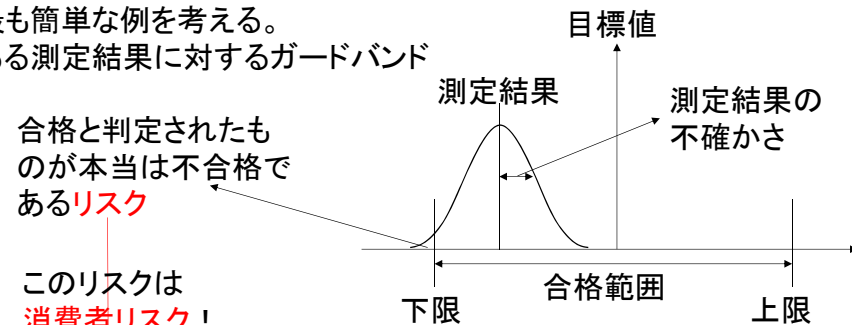
不確かさを取り入れた場合いろいろなパターンが考えられる。



当事者間の合意がない場合には, このみが合格。
(ISO14253-1, JIS B 0641-1の場合)

リスクによるガードバンドの設定

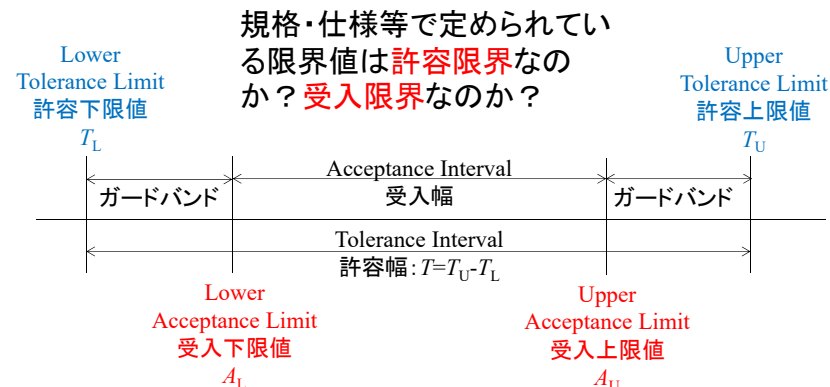
最も簡単な例を考える。
ある測定結果に対するガードバンド



このリスクを設定し, ガードバンドを決定する。
例えば, 消費者リスク2.5%まで許容されるとすると, ガードバンドは4度拡張不確かさと等しくなる。

ある一つの測定結果に対するリスクのことを“Specific Risk”という!

適切なガードバンドの導入には?

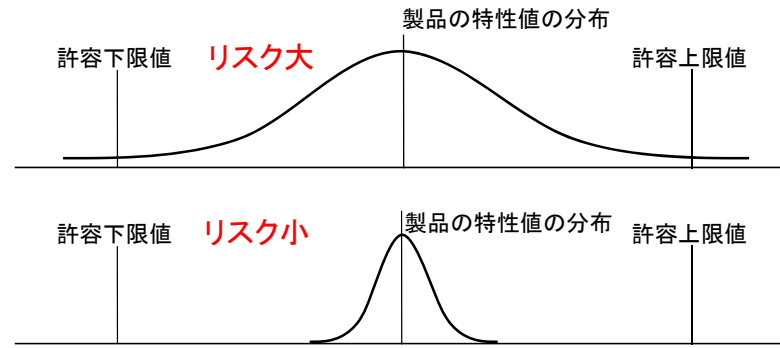


規格・仕様等で定められている限界値がすでに安全を見て, 受入限界値を定めている場合がある! このような場合に更にガードバンドを設定するのは厳しすぎる限界値を定めることとなる。

更なる拡張

これまでは、測定結果の不確かさのみを考え、合否判定のリスクを考慮した。

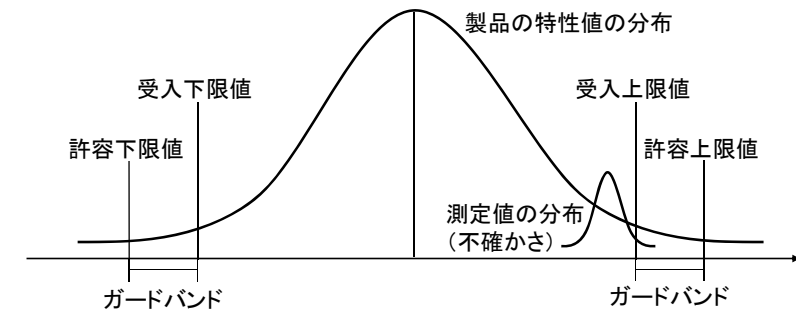
更に考えを進めると、合否判定を行う製品が複数ある場合には、製品の特性値の分布も存在するはずであり、特性値の分布もリスクに直結するはずである。



更なる拡張

ここではILAC G8の参考文献にあるDobbertの方法を紹介する。

合否判定を行う製品が複数ある場合には、製品の特性値の確率分布と測定値の確率分布の両者を重ね合わせてリスクを計算することができるはず。このようなリスクを“Global Risk”という。



2つのリスク

3.3.13

specific consumer's risk

probability that a **particular** accepted item is non-conforming

3.3.15

global consumer's risk

consumer's risk

probability that a non-conforming item will be accepted based on a **future measurement result**

Specificのほうは、何かを測定した結果があって、その結果のリスクのこと。Globalのほうは、将来の結果について、つまりすでに得られている結果ではなく、そのアイテム全体で考えられるリスクのこと。

更なる拡張

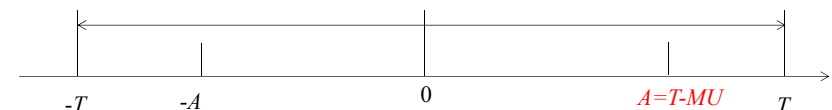
例として消費者リスクを考える

$$P_{CUS} = \int_{-T}^{\infty} \int_{-A}^A p_0(e_{dut}) p_m(y - e_{dut}) dy de_{dut} + \int_{-\infty}^{-T} \int_{-A}^A p_0(e_{dut}) p_m(y - e_{dut}) dy de_{dut}$$

ここで、既知のものと未知のものを分類。

- ・既知: 許容値 T , 測定の合成標準不確かさ σ_m
- ・未知: 確率密度関数の p_0 , p_m
- ・求めたいもの: $A = T - MU$

ここで、 $U = 2\sigma_m$ であり ($U = 1.96\sigma_m$ とする場合もあり。), M は、Guardband Factor と呼ばれる、ガードバンドの大きさを決定する値で、拡張不確かさの何倍をガードバンドとするのかを示す係数である。ここで、未知のものを前提を置いて設定する。



更なる拡張

$$P_{\text{CUS}} = \int_T^\infty \int_{-A}^A p_0(e_{\text{dut}}) p_m(y - e_{\text{dut}}) dy de_{\text{dut}} + \int_{-\infty}^{-T} \int_{-A}^A p_0(e_{\text{dut}}) p_m(y - e_{\text{dut}}) dy de_{\text{dut}}$$

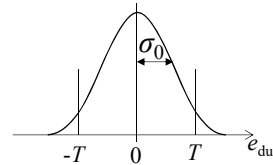
製品の特性値の分布は正規分布に従っているとする。
 これまで行ってきた試験において特性値が $\pm T$ 以内に入っていた確率 p を調べる。
 それが分かれば、製品の特性値の標準偏差は、

$$\sigma_0 = \frac{T}{\text{NORMSINV}\left(\frac{1-p}{2}\right)}$$

となり、製品の特性値の分布は、

$$e_{\text{dut}} \sim N(0, \sigma_0^2)$$

となる。



更なる拡張

$$P_{\text{CUS}} = \int_T^\infty \int_{-A}^A p_0(e_{\text{dut}}) p_m(y - e_{\text{dut}}) dy de_{\text{dut}} + \int_{-\infty}^{-T} \int_{-A}^A p_0(e_{\text{dut}}) p_m(y - e_{\text{dut}}) dy de_{\text{dut}}$$

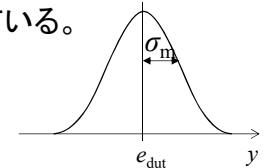
測定値の分布は正規分布に従っているとする。
 合成標準不確かさが既知であり、 σ_m で表されている。
 よって、測定値の分布は、

$$y - e_{\text{dut}} \sim N(0, \sigma_m^2)$$

となる。

また多くの場合、許容差 T と拡張不確かさ $U=2\sigma_m$ との比である、Test Uncertainty Ratio(TUR)が求められている。

$$TUR = T/U$$

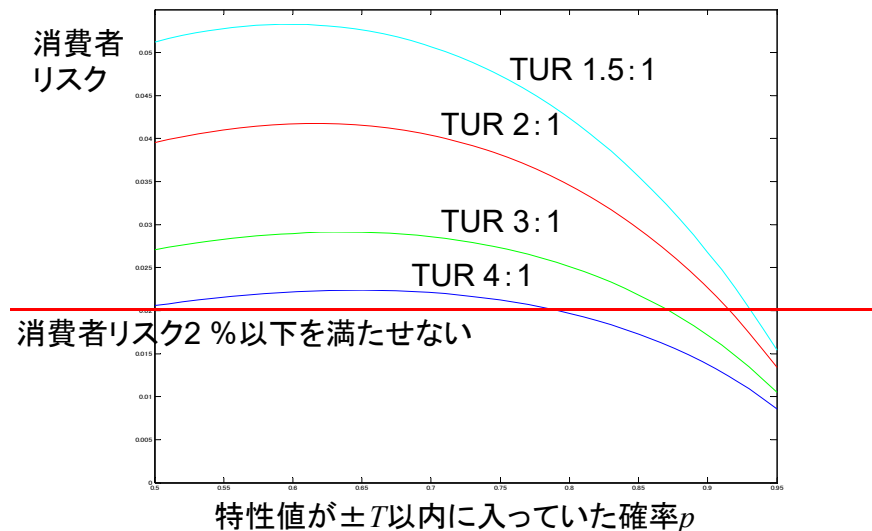


更なる拡張

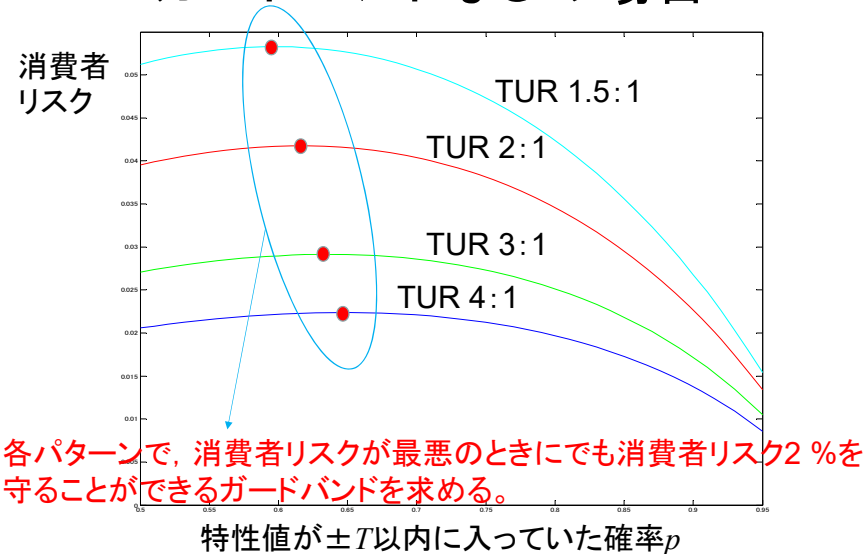
$$P_{\text{CUS}} = \int_T^\infty \int_{-A}^A p_0(e_{\text{dut}}) p_m(y - e_{\text{dut}}) dy de_{\text{dut}} + \int_{-\infty}^{-T} \int_{-A}^A p_0(e_{\text{dut}}) p_m(y - e_{\text{dut}}) dy de_{\text{dut}}$$

そうすると、上記の式は、
 これまで行ってきた試験において特性値が $\pm T$ 以内に入っている確率 p と、
 許容差 T と拡張不確かさ $U=2\sigma_m$ との比である、Test Uncertainty Ratio(TUR)を与えれば解くことができ、リスクが計算できる。

ガードバンドなしの場合



ガードバンドなしの場合



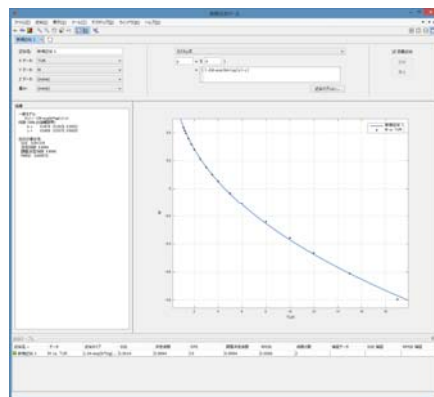
フィッティング

TUR	In Tolerance p	Max Risk	M × 100%
1.1	57.15%	6.956%	43.68%
1.2	57.89%	6.495%	41.56%
1.3	58.54%	6.092%	39.58%
1.5	59.62%	5.420%	35.90%
1.75	60.67%	4.763%	31.74%
2	61.50%	4.249%	27.95%
2.5	62.71%	3.495%	21.24%
3	63.55%	2.968%	15.37%
3.5	64.18%	2.579%	10.13%
4	64.65%	2.281%	5.33%
5	65.34%	1.852%	-3.22%
6	65.80%	1.559%	-10.81%
8	66.40%	1.184%	-24.07%
10	66.76%	0.955%	-35.72%
12	67.01%	0.800%	-46.36%
15	67.26%	0.643%	-61.12%
19	67.47%	0.510%	-79.48%

注：原文と若干異なる数値があります。

x軸をTUR, y軸をMとして関数にフィッティング。

フィッティング



多少値は異なっているが、大体論文と同じ値が出る。

つまり、Dobbertの方法は、消費者リスク2 %を守るためのもの。それ以外のときには使えない。

$$M_{2\%} = 1.04 - \exp[0.38 \log(TUR) - 0.54]$$

つまり、TURさえ分かれば、上式でガードバンドを設定でき、その結果最悪でも消費者リスク2 %は守れる。

Deaverの方法

消費者リスク2 %を守るためにDobbertの方法が優れている。でも、Dobbertの数式みたいな面倒なものを使わなくても、

$$A = \sqrt{T^2 - U^2} = \sqrt{T^2 - \left(\frac{T}{TUR}\right)^2} = T \sqrt{1 - \frac{1}{TUR^2}}$$

で、よくないか？上記の値のことを二乗したものの差の平方根なので、Root-Difference-of-Squares(RDS)と呼ぶ。

2法を比較

$$K_{\text{Dobbert}} = \frac{A}{T} = \frac{T - MU}{T} = \frac{T - (1.04 - \exp[0.38 \log(TUR) - 0.54]) \frac{T}{TUR}}{T} = 1 - \frac{1.04 - \exp[0.38 \log(TUR) - 0.54]}{TUR}$$

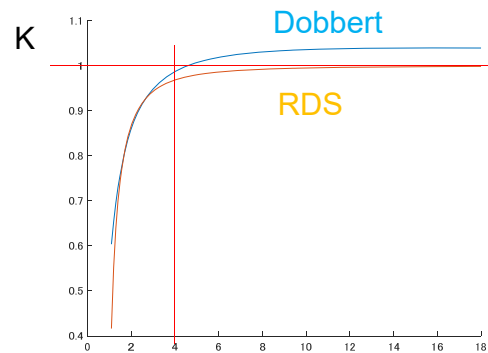
$$K_{\text{RDS}} = \frac{A}{T} = \sqrt{1 - \frac{1}{TUR^2}}$$

両者ともAとTの比であるKを求めて比較する。

注：Deaverの方法と呼ばれるものは複数あります。ここで紹介したのは、ILAC G8で引用されているDeaverの論文の方法です。

比較結果

TUR	1.1	1.5	2	2.5	3	3.5	3.9	4
Dobbert	0.60	0.76	0.86	0.91	0.95	0.97	0.98	0.99
RDS	0.42	0.75	0.87	0.92	0.94	0.96	0.97	0.97



ちなみに $K=4$ 以上を問題にしないのは、4:1ルールのほうが満たせるから。

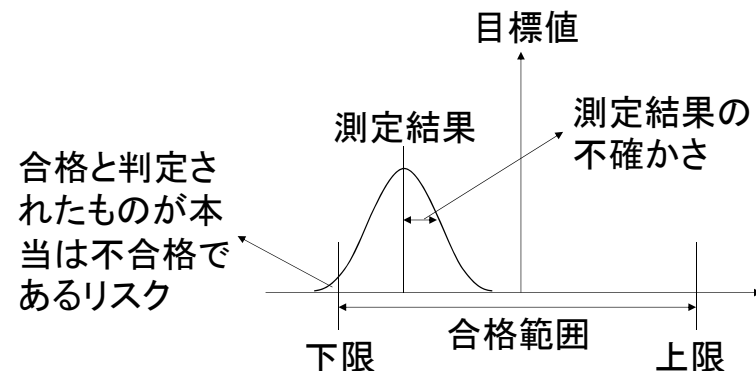
ガードバンドが必要なくなる $A/T=1$ あたりまで十分に近い値である。

Deaverの論文の結論いろいろの方法はあるけど、面倒くさかったり、大きなガードバンドになりすぎたりするので、簡単で妥当なRDSがおすすめ。

合否判定とガードバンド

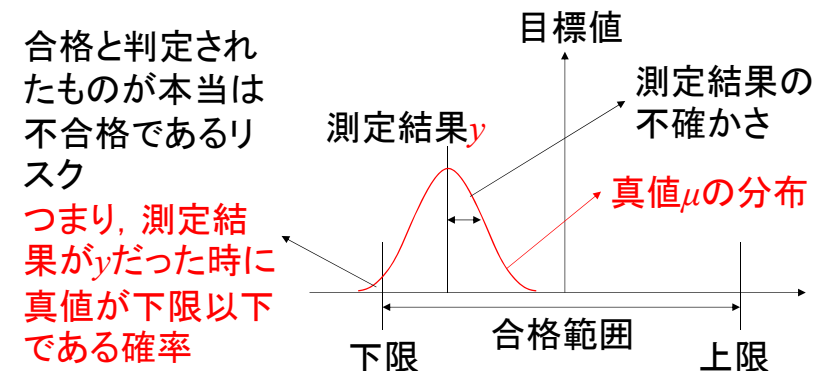
- 顧客との合意はどうなっていますか？
- ガードバンドを本当に設定する必要がありますか？
- 考慮すべきリスクはSpecific Riskですか？Global Riskですか？
- Specific Riskの場合は正規分布を基にして簡単にガードバンドが決定できるでしょう。
- Global Riskの場合、DobbertやDeaverの方法は、消費者リスク2%を基にしたものでさらに簡単な方法ですが、消費者リスク2%というのは妥当ですか？この2%はあくまでもANSI/NCSL Z540.3で決められているものです。
- Global Riskをちゃんと計算するためには、ある程度の数値計算が必須になります。チャートを利用する方法もあります。

おまけ: ベイズ統計



おまけ: ベイズ統計

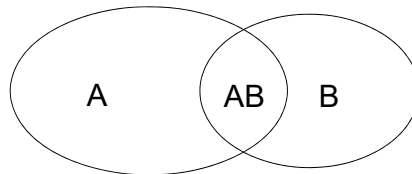
実はこの考え方はベイズ統計！
つまり、真値が変数であり、真値の分布を考える。



つまり、測定結果が y だった時に真値が下限以下である確率

おまけ: ベイズの定理

事象



事象Aが起こる確率 $p(A)$

事象Bが起こる確率 $p(B)$

事象AとBが両方起こる確率 $p(AB)$

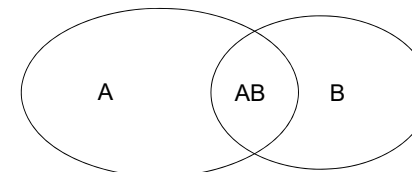
事象Aが起こっていたとして、事象Bが起こる確率 $p(B|A)$

事象Bが起こっていたとして、事象Aが起こる確率 $p(A|B)$

とすると、

おまけ: ベイズの定理

事象



$$p(AB) = p(A)p(B|A) = p(B)p(A|B)$$

が成立する。ここで、

$$p(A)p(B|A) = p(B)p(A|B)$$

$$p(B|A) = \frac{p(A|B)}{p(A)} p(B)$$

これがベイズの定理である。

言い換えると、最初Bについての事前の情報 $p(B)$ を持っていたが、新たにAというBに関する知識を得たので、事前の情報が更新されて、新しい情報、 $p(B|A)$ を得た、ということ。

おまけ: ベイズ統計

確率密度関数で書き表すと、

$$f(q|d) = Kl(q|d) f(q)$$

事後pdf (posterior)

事前pdf (prior)

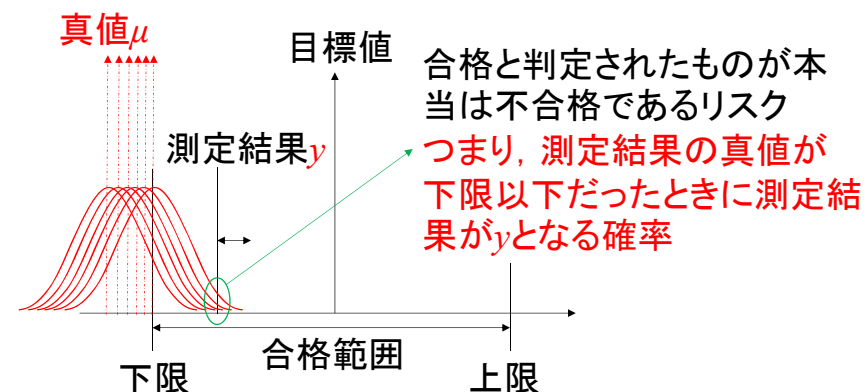
となる。言い換えると、未知母数の事前pdfをデータから求められる関数によって変形し、事後pdfが得られる、ということである。

つまり、母数に変数となっている！

おまけ: ベイズ統計

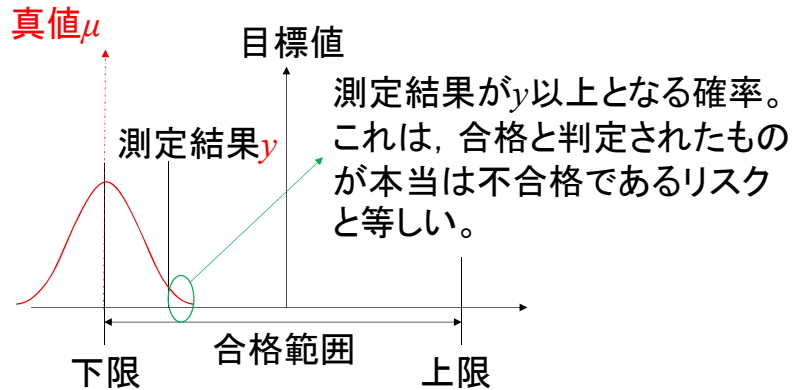
これまでの統計(頻度統計)だと？

つまり真値は定数であり、測定結果が変数。



おまけ: ベイズ統計

頻度統計の場合では下の図で説明されることが多いが、実際に図で表されている確率は、真値が下限と一致した場合、測定結果が y 以上となる確率。



参考文献

- ILAC-G8:09/2019, Guidelines on Decision Rules and Statements of Conformity.
- OIML G 19/2017, The role of measurement uncertainty in conformity assessment decisions in legal metrology.
- JCGM 106:2012, Evaluation of measurement data - The role of measurement uncertainty in conformity assessment.

参考文献

- ANSI/NCSL Z540.3-R2013, Requirements for the Calibration of Measuring and Test Equipment.
- Dobbert, M., "A Guard-Band Strategy for Managing False-Accept Risk", Proc., NCSL Workshop & Symposium, 2008.
- Deaver, D, and Somppi, J., "A study of and recommendation for applying the false acceptance risk specification of Z540.3", Proc., NCSL Workshop & Symposium, 2007.