

確率統計序説

秋田 智之

平成 20 年 6 月 23 日

概 要

このノートは工科系数学統一試験の確率統計の出題範囲で挙げられた項目を整理し、2006 年までの過去問を並べたものです。問題では には適当な文字・数字を、【 】は選択肢の中から適切なものを選んでください。本稿に収めた内容のみではとても確率統計の必要な事柄を全て網羅できていませんが、勉強の第一歩には使えると思います。

1 確率の基礎観念

1.1 確率と事象の独立性

この節ではサイコロ投げやコイン投げを通じて事象や確率を解説する。以下では Ω を**標本空間**とする。これは起こりうる全ての現象を集めたものである。また標本空間の元を**標本**という。

例 コインを 1 回投げる試行においては $\Omega = \{\text{表}, \text{裏}\}$ 。サイコロを 1 回投げる試行においては $\Omega = \{1 \text{ の目が出る}, 2 \text{ の目が出る}, \dots, 6 \text{ の目が出る}\}$ となる。また、コインを 2 回投げる試行においては $\Omega = \{(\text{表}, \text{表}), (\text{表}, \text{裏}), (\text{裏}, \text{表}), (\text{裏}, \text{裏})\}$ となる。

早速確率を定義したいのだが、その前に確率が定義できる**事象**を考えよう。事象とは Ω の部分集合である。

注意。 正確には事象を定義するにはもう少し条件があります。

例 サイコロ投げにおいて $A = \{1, 3\}$ という Ω の部分集合を考える。これは「1 または 3 の目が出るという」事象を表している。

ここで**確率**を定義する。確率とは事象 A に対して、 A の起こりやすさを 0 から 1 までの値で表したものである。

注意。 正確には確率を定義するときにももう少し条件が必要です。

例 正確に作られたコインを1回投げる試行では

$$P(\phi) = 0, \quad P(\{\text{表}\}) = \frac{1}{2}, \quad P(\{\text{裏}\}) = \frac{1}{2}, \quad P(\Omega) = 1$$

となります。正確に作られたサイコロを1回投げる試行においては

$$P(A) = \frac{1}{6} \times (A \text{ の元の個数})$$

となります。

問 (2004 年統一試験). A, B を事象, A^c, B^c をそれぞれ A, B の余事象とする. $P(A) = \frac{1}{2}, P(B) = \frac{2}{3}, P(A \cap B) = \frac{1}{3}$ のとき, $P(A^c \cap B^c) = \frac{\square}{\square}$ である.

次に条件付確率を定義します。次の例を見てください。

例 正確に作られたサイコロを投げる試行において事象 A を「奇数の目が出る」、事象 B を「3または6の目が出る」とします。このとき A が起こったという条件下で B が起こる確率を考えてみましょう。まず事象 A が起こったのでサイコロの目は1か3か5のいずれかです。このとき事象 B が起こる、つまり3の目が出る確率は $1/3$ となります。

高校までの定義によると、上の条件付確率は $\sharp(A) = 3, \sharp(A \cap B) = 1$ であることを用いて $\sharp(A \cap B)/\sharp(A)$ によって計算されています。この式の分子・分母を $\sharp(\Omega)$ で割ると

$$\frac{\sharp(A \cap B)/\sharp(\Omega)}{\sharp(A)/\sharp(\Omega)} = \frac{P(A \cap B)}{P(A)}$$

となります。この右辺に現れる式を条件付確率の定義に採用します。

定義 (条件付確率) .

A, B を事象とし, $P(B) \neq 0$ とする. このとき B の下での A の条件付確率 $P(A|B)$ を次で定義する.

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$P(A|B)$ は $P_B(A)$ と書かれることもある。

ここで

$$P(A|B) = P(A)$$

となっている場合を考えよう。これは A が起きる確率が「 B が起こった元で A が起きる確率」と等しいことを表しています。つまり B が起きようが起きまいが A の起こる確率が変

化しないということになります。このとき A と B は独立であるといいます。条件付確率の定義より、次を事象が独立であることの定義にします。

定義 (独立と従属) .

事象 A, B について $P(A \cap B) = P(A)P(B)$ が成り立つとき、 A と B は**独立**であるという。そうでないときは**従属**であるという。

例 正確に作られたサイコロを投げる試行において $A = \{1, 3, 5\}$, $B = \{4, 5\}$ とすると $P(A) = 1/2$, $P(B) = 1/3$, また $A \cap B = \{5\}$ なので $P(A \cap B) = 1/6$ となっていて $P(A)P(B) = 1/6 = P(A \cap B)$ が成り立っているので A と B は独立です。一方 $C = \{1, 3\}$ とすると $P(C) = 1/3$ であり $A \cap C = \{1, 3\}$ なので $P(A \cap C) = 1/3$ となり, $P(A)P(C) = 1/6 \neq 1/3 = P(A \cap C)$ なので A と C は従属となります。

問 (2004 年統一試験). 確率が $P(A) \neq 0$, $P(B) \neq 0$ である 2 つの事象 A, B が $A \cap B = \phi$ を満たすならば A と B は独立である。ただし, ϕ は空集合とする。【Yes, No】

1.2 確率変数と分布

例えば 1 回コインを投げる試行において表を 1, 裏を 0 と対応させると数学的に扱いやすくなります。このように標本を数値化したものを**確率変数**といいます。

注意. 正確には全ての実数 x について「確率変数 X が x 以下をとるという確率」が定義されている必要があります。また, Ω が数集合のときは標本と確率変数は同一視されることがあります。

例 サイコロを 1 回投げる試行においては「出た目の数」などが確率変数になります。またコインを 2 回投げる試行においては「表が出た回数」などが確率変数になります。

確率変数が定義できると次の分布関数を定義することができます。

定義 (分布関数) .

X を確率変数とする。このとき確率変数 X が x 以下であるという確率

$$F(x) = P(X \leq x)$$

を X の**累積分布関数**または単に**分布関数**という。分布関数は右連続な単調増加関数になる。

例 離散的な確率変数 X の分布が

$$P(X=1) = \frac{1}{3}, \quad P(X=2) = \frac{2}{3}$$

であるときの分布関数を求めます. $x \geq 2$ のとき $X \leq x$ を満たす X は 1 と 2 ですから $F(x) = P(X \leq x) = p_1 + p_2 = 1$ となります. $1 \leq x < 2$ のとき $X \leq x$ を満たす X は 1 のみですから $F(x) = p_1 = 1/3$ です. $x < 1$ のときは $X \leq x$ を満たすものはありませんので $F(x) = 0$ です. これらをまとめると

$$F(x) = \begin{cases} 0 & (x < 1) \\ 1/3 & (1 \leq x < 2) \\ 1 & (x \geq 2) \end{cases}$$

となります. (不等号が $<$ なのか $\leq$ なのかは (離散型の場合は) きちんと区別してください.)

定理 (2005 年統一試験). 確率変数 X の分布関数を $F_X(x) = P(X \leq x)$ とする. このとき $\lim_{x \rightarrow -\infty} F_X(x) = \square$, $\lim_{x \rightarrow \infty} F_X(x) = \square$ である.

問 (2006 年統一試験). 確率変数 X の分布関数を $F(x) = P(X \leq x)$ とする. このとき $F(0) = 0$, $F(2) = 1$ ならば $F(-1) = \square$ であり $P(X > 3) = \square$ である.

連続型の確率変数の場合, 分布関数はある非負の関数 $f(x)$ の積分で表現されます.

定義 (密度関数). —

連続型確率変数 X の分布関数 $F(x)$ が, ある関数 $f(x)$ を用いて

$$F(x) = \int_{-\infty}^x f(t) dt$$

と書けるとき, $f(x)$ を X の **確率密度関数** または単に **密度関数** という.

定理. $f(x)$ を密度関数とすると $\int_{-\infty}^{\infty} f(x) dx = 1$ が成り立つ.

問 (2004 年統一試験). 確率変数 X の確率密度関数 $f(x)$ がある定数 C に対して

$$f(x) = \begin{cases} Ce^{2x} & x \leq 0 \text{ のとき} \\ 0 & x > 0 \text{ のとき} \end{cases}$$

で与えられているとき、 $C = \boxed{}$ である。したがって、 X の値が x 以下になる確率として定義される分布関数 $F(x) = P(X \leq x)$ は $x \leq 0$ のとき $\boxed{}$ であり、これから、 $P(X > -\log 2) = \boxed{}$ が分かる。一方 $x > 0$ のとき $F(x) = \boxed{}$ である。

問 (2003年統一試験). 確率変数 X の確率密度関数は次で与えられる。

$$f(x) = \begin{cases} \frac{1}{2} & (-1 \leq x \leq 1) \\ 0 & (\text{その他}) \end{cases}$$

このとき、 X の分布関数 $F_X(x)$ は

$$F_X(x) = \int_{-\infty}^x f(s) ds$$

であるから、 $-\infty < x < -1$ のときは $F_X(x) = \boxed{}$ である。次に $-1 \leq x \leq 1$ のときは

$$F_X(x) = \int_{-1}^x \boxed{} ds$$

となるから、この積分を計算して $F_X(x) = \boxed{}$ を得る。更に $1 < x$ のときは $F_X(x) = \boxed{}$ となる。

定義 (確率関数). —

X を離散型確率変数とする。このとき X がある値 x をとる確率

$$p(x) = P(X = x)$$

を**確率関数**という。

定理. $p(x)$ を確率関数とすると $\sum_x p(x) = 1$ が成り立つ。($\sum_x$ はとりうる全ての x についての $p(x)$ の和であると考えことにする.)

例 (正確に作られた)サイコロ投げでは

$$p(x) = \begin{cases} 1/6 & (x = 1, 2, 3, 4, 5, 6) \\ 0 & (\text{それ以外}) \end{cases}$$

となります。

次に確率変数が複数ある場合を考えよう.

定義 (同時確率関数).

X, Y が離散型確率変数であるとき

$$p(x, y) = P(X = x, Y = y)$$

を X, Y の**同時確率関数**という.

注意. 連続型の場合も同様に同時確率密度関数というものが考えられます.

確率変数 X, Y が独立であることの定義も事象の独立性と同様に次で定義されます.

定義 (確率変数の独立性).

離散型確率変数 X, Y の確率関数をそれぞれ $p_X(x), p_Y(y)$, X, Y の同時確率関数を $p(x, y)$ とする. このとき全ての x, y について

$$p(x, y) = p_X(x)p_Y(y)$$

が成り立つとき, X, Y は**独立**であるという. また, 一組の x, y についても成り立たないものがあれば x, y は**従属**であるという.

注意. 連続型の場合も同様に独立性は定義できます.

例 正確なサイコロを投げる試行において

$$X = \begin{cases} -2 & (\text{出た目が偶数}) \\ 2 & (\text{出た目が奇数}) \end{cases} \quad Y = \begin{cases} -1 & (\text{出た目が } 1, 2, 3) \\ 0 & (\text{出た目が } 4) \\ 1 & (\text{出た目が } 5, 6) \end{cases}$$

とする. このとき X と Y が独立か従属かを調べてみよう. 上の確率変数の独立性の定義から全ての x, y について

$$P(X = x, Y = y) = P(X = x)P(Y = y) \quad (\spadesuit)$$

が成り立っていれば独立, そうでなければ従属である. まず, $X = -2, Y = -1$ となる場合を調べよう. $X = -2$ かつ $Y = -1$ となるのは出た目が 2 のときのみであるから $P(X = -2, Y = -1) = 1/6$ である. 一方, $P(X = -2) = 1/2$ で, $P(Y = -1) = 1/2$ であるので, $P(X = -2)P(Y = -1) = 1/4 \neq 1/6 = P(X = -2, Y = -1)$ である. よって $(\spadesuit)$ を満たさない例があるので X と Y は従属である.

1.3 期待値 (平均) と分散

この節では確率分布の特徴を表す量である期待値や分散について扱う。

定義 (期待値と分散) .

確率変数 X の**期待値** $E(X)$ を次で定義する. ($E(X)$ を μ で表すことも多い.)

$$E(X) = \begin{cases} \sum x \cdot p(x) & (\text{離散型のとき}) \\ \int_{-\infty}^{\infty} x f(x) dx & (\text{連続型のとき}) \end{cases}$$

ここで $p(x)$ は離散型確率変数の確率関数で, $f(x)$ は連続型確率変数の密度関数である. これを一般化して $g(x)$ を関数としたとき, **期待値** $E(g(X))$ を次で定義する.

$$E(g(X)) = \begin{cases} \sum g(x) \cdot p(x) & (\text{離散型のとき}) \\ \int_{-\infty}^{\infty} g(x) f(x) dx & (\text{連続型のとき}) \end{cases}$$

特に $g(x) = (x - \mu)^2$ としたときの $E((X - \mu)^2)$ を**分散**といい, $V(X)$ や $\text{Var}(X)$ で表す.

定理 (2005 年統一試験). 確率変数 X に対し, $E(X)$ を X の期待値とする. このとき

$$E(\{X - E(X)\}^2) = E\left(\boxed{}\right) - \left\{E\left(\boxed{}\right)\right\}^2$$

である. (分散 = 2 乗の平均 - 平均の 2 乗)

問 (サンプル問題). 分布の中心を表すのが母平均 (期待値) であるから, 分布が与えられれば必ず母平均は存在する. 【Yes, No】

問 (サンプル問題). 母平均が存在しなくても母分散が存在する場合がある. 【Yes, No】

問 (2006 年統一試験). 離散型の確率変数 Y の確率分布は

Y の値	0	1	2	3
確率	0.6	0.1	0.2	0.1

で与えられるとする. Y の期待値は $E(Y) = \boxed{}$ であり, 分散は $V(Y) = \boxed{}$ である.

問 (2004 年統一試験). 確率変数 X の期待値 $E(X)$ が 2 であっても $E\left(\frac{1}{X}\right) = \frac{1}{2}$ であるとは限らない. 【Yes, No】

定理. X, Y を確率変数, c を定数とするとき

$$E(X + Y) = E(X) + E(Y), \quad E(cX) = cE(X), \quad V(cX) = c^2V(X)$$

が成り立つ. また, X と Y が**独立ならば**

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

が成り立つ.

注意. X と Y が独立でないときでも

$$V(X + Y) = V(X) + V(Y) + 2\text{Cov}(X, Y)$$

は成り立ちます.

確率変数 X, Y の関係を表す量として共分散や相関係数があります.

定義 (共分散, 相関係数). —

X, Y を確率変数とし, $E(X) = \mu_X, E(Y) = \mu_Y$ とする. このとき

$$\text{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)] = \begin{cases} \sum (x - \mu_X)(y - \mu_Y)p(x, y) & (\text{離散型}) \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_Y)f(x, y)dx dy & (\text{連続型}) \end{cases}$$

を X と Y の**共分散**という. また

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{V(X)}\sqrt{V(Y)}}$$

を X と Y の**相関係数**という.

注意. X が大きい値をとれば Y も大きな値をとり, X が小さな値をとれば Y も小さな値をとる傾向があるとき相関係数は 1 に近くなります. X が大きい値をとれば Y は小さな値をとり, X が小さな値をとれば Y は大きな値をとる傾向があるとき相関係数は -1 に近くなります.

定理. 確率変数 X, Y が独立ならば X と Y との相関係数は 0 である.

定理. $-1 \leq \rho(X, Y) \leq 1$ が成り立つ.

1.4 代表的な確率分布

この節では代表的な離散型確率分布である二項分布，ポアソン分布，代表的な連続型確率分布である一様分布，指数分布，正規分布を扱う．分布の導き方も一応は載せたが難しいところもあるのでそこは読み飛ばしても構わない．

例 画鋲を投げて「上」となる確率を p ，それ以外になる確率を $1-p$ とします．この画鋲を 4 回投げて 2 回表が出る確率を求めます．まず，4 回投げて 2 回表が出る事象は「上，上，他，他」「上，他，上，他」「上，他，他，上」「他，上，上，他」「他，上，他，上」「他，他，上，上」の 6 通りです．これは 4 つの中から 2 個を選ぶ組み合わせの個数 ${}_4C_2$ に一致します．次に各標本から得られる確率を求めると全て $p^2(1-p)^2$ となります．従って，4 回中 2 回表が出る確率はこれらの積

$${}_4C_2 p^2 (1-p)^2$$

となります．

定義&定理 (二項分布) .

p は $0 \leq p \leq 1$ を満たす定数とする．離散型確率変数 X の確率関数が

$$p(x) = {}_nC_x p^x (1-p)^{n-x} \quad (x = 0, 1, \dots, n)$$

で表されているとき X の従う分布をパラメータ n, p の**二項分布**といい， $B(n, p)$ で表す．

定理． X が二項分布 $B(n, p)$ に従うとき， $E(X) = np, V(X) = np(1-p)$ である．

問 (2005 年統一試験)．確率変数 X, Y が独立で同一の二項分布 $B(5, \frac{2}{3})$ に従うものとするとき，条件付き確率 $P(Y = 1 | X + Y = 2)$ を求めたい． X, Y は二項分布 $B(5, \frac{2}{3})$ に従うから確率関数 $p(k) = P(X = k) = P(Y = k)$ は

$$p(k) = \square C_{\square} \left(\frac{2}{3}\right)^{\square} \left(\frac{1}{3}\right)^{\square}, \quad k = 0, 1, \dots, 5$$

である． $P(Y = 1 | X + Y = 2)$ は条件 $X + Y = 2$ の下で $Y = 1$ となる確率である． X, Y はともに 0 以上の整数値をとるから，条件 $X + Y = 2$ を満たすのは $(X, Y) = (0, 2), (1, 1), (2, 0)$ の 3 つの場合のみである． X, Y は独立であるから

$$P(X = 0, Y = 2) = p(0)p(2) = \frac{\square}{\square}, \quad P(X = 1, Y = 1) = \frac{\square}{\square}, \quad P(X = 2, Y = 0) = \frac{\square}{\square}$$

となる．このうち $Y = 1$ となるのは $(X, Y) = (1, 1)$ のときのみである．したがって，

$$P(Y = 1 | X + Y = 2) = \frac{\square}{\square}$$

となる.

問 (2004 年統一試験). n は自然数, p は $0 < p < 1$ を満たす定数であるとする. 確率変数 $X_1, X_2, \dots, X_n$ は独立で, 各 $k = 1, 2, \dots, n$ について X_k が値 $1, 0$ をとる確率はそれぞれ

$$P(X_k = 1) = p, \quad P(X_k = 0) = 1 - p$$

であるとする. このとき X_k の期待値 $E(X_k)$ は , X_k^2 の期待値 $E(X_k^2)$ は であり, X_k の分散 $V(X_k)$ は である.

次に

$$S = \sum_{k=1}^n X_k$$

とおくと, その期待値 $E(S)$ は である. また $X_1, X_2, \dots, X_n$ が独立であることに注意すれば, 分散 $V(S)$ は となる. ところで, S は $X_1, X_2, \dots, X_n$ のうち値 1 をとるものの個数を表す確率変数で 分布に従う. 特に, $p = \frac{1}{2}$, $n = 5$ のとき, $P(S = 3) = \frac{\text{}}{\text{}}$ である.

次に二項分布において n が大きく p がすごく小さい場合を考えます. 例えば沢山の製品のうち不良品が m 個以下である確率などを考えることを想定しています. この場合, 二項分布の式から $P(X = 0) + P(X = 1) + \dots$ と計算していくのは大変です. そこで $np = \lambda$ として次の変形を考えます.

$$\begin{aligned} P(X = x) &= \frac{n!}{x!(n-x)!} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} \\ &= \frac{\lambda^x}{x!} \frac{n(n-1)\cdots(n-x+1)}{n^x} \left(1 - \frac{\lambda}{n}\right)^{-x} \left(1 - \frac{\lambda}{n}\right)^n \end{aligned}$$

ここで極限を考えたいのですが, それぞれ $n \rightarrow \infty$, $p \rightarrow 0$ としても $P(X = x)$ は収束しないので期待値 $np = \lambda$ を固定して $n \rightarrow \infty$, $p \rightarrow 0$ とすると $n/n, (n-1)/n, \dots, (n-k+1)/n, (1 - \lambda/n)^{-x}$ は全て 1 に収束し, $(1 - \lambda/n)^n$ は $e^{-\lambda}$ に収束するので次が得られます.

定義&定理 (ポアソン分布).

λ を正の定数とする. 離散型確率変数 X の確率関数が

$$p(x) = \frac{\lambda^x}{x!} e^{-\lambda} \quad (x = 0, 1, 2, \dots)$$

で表されているとき X の従う分布をパラメータ λ の**ポアソン分布**という.

定理. X がパラメータ λ のポアソン分布に従うとき $E(X) = \lambda$, $V(X) = \lambda$ となる.

例 ポアソン分布の期待値 $E(X)$ は次のように計算できます.

$$\begin{aligned} E(X) &= \sum_{x=0}^{\infty} x e^{-\lambda} \frac{\lambda^x}{x!} = \sum_{x=1}^{\infty} x e^{-\lambda} \frac{\lambda^x}{x!} = e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^x}{(x-1)!} \\ &= \lambda e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^{(x-1)}}{(x-1)!} = \lambda e^{-\lambda} \sum_{y=0}^{\infty} \frac{\lambda^y}{y!} = \lambda e^{-\lambda} e^{\lambda} = \lambda \end{aligned}$$

分散 $V(X)$ を計算するために $E(X^2)$ を求めたいのですが, この場合は直接 $E(X^2)$ を計算するよりも $E(X(X-1))$ を求めて $E(X^2) = E(X(X-1)) + E(X)$ を利用したほうが簡単です.

$$\begin{aligned} E(X(X-1)) &= \sum_{x=0}^{\infty} x(x-1) e^{-\lambda} \frac{\lambda^x}{x!} = \sum_{x=2}^{\infty} x(x-1) e^{-\lambda} \frac{\lambda^x}{x!} = e^{-\lambda} \sum_{x=2}^{\infty} \frac{\lambda^x}{(x-2)!} \\ &= \lambda^2 e^{-\lambda} \sum_{x=2}^{\infty} \frac{\lambda^{(x-2)}}{(x-2)!} = \lambda^2 e^{-\lambda} \sum_{y=0}^{\infty} \frac{\lambda^y}{y!} = \lambda^2 e^{-\lambda} e^{\lambda} = \lambda^2 \end{aligned}$$

従って $E(X^2) = \lambda^2 + \lambda$ なので $V(X) = E(X^2) - (E(X))^2 = \lambda^2 + \lambda - (\lambda)^2 = \lambda$ を得ます.

定義 (一様分布).

a, b を $a < b$ を満たす定数とする. このとき連続型確率変数 X の密度関数が

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

で表されているとき X は $[a, b]$ 上の**一様分布**に従うという.

問 (2006 年統一試験). c は $c > -4$ を満たす定数とし, 確率変数 X は閉区間 $[-4, c] = \{x \mid -4 \leq x \leq c\}$ 上の一様分布に従っているものとする. このとき $P(X \leq -2) = \frac{1}{3}$ であるとする, $c = \boxed{}$ である. また期待値は $E(X) = \boxed{}$ であり, $P(-3 \leq X \leq -1) = \boxed{}$ である.

次に「蛍光灯が切れるまでの時間」や「ある機械が故障するまでの時間」のように寿命を表す確率分布を考えます. 簡単のため「ある時刻 t まで切れなかった蛍光灯が次の瞬間に切れる確率は常に一定で λ である」という仮定をします. このとき X を蛍光灯が切れるまでの時間を表す確率変数, $F(x)$ をその分布関数とすると

$$\begin{aligned} \lambda &= \lim_{h \rightarrow 0} \frac{1}{h} P(t \leq X \leq t+h \mid X \geq t) = \lim_{h \rightarrow 0} \frac{1}{h} \frac{P((t \leq X \leq t+h) \cap (X \geq t))}{P(X \geq t)} \\ &= \lim_{h \rightarrow 0} \frac{1}{h} \frac{P(t \leq X \leq t+h)}{1 - F(t)} = \lim_{h \rightarrow 0} \frac{1}{h} \frac{(F(t+h) - F(t))}{1 - F(t)} = \frac{F'(t)}{1 - F(t)} \end{aligned}$$

となります。この微分方程式は $G(x) = 1 - F(x)$ とおくと変数分離型になり次のように解が得られます。

$$\lambda = \frac{G'(x)}{G(x)} \quad \therefore G(x) = Ce^{-\lambda x}$$

これを $F(x)$ に戻し、微分して $\int_{-\infty}^{\infty} f(x)dx = 1$ を満たすように定数 C を求めると次が得られます。

定義 (指数分布) .

λ を正の定数とする。このとき連続型確率変数 X の密度関数が

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & (x \geq 0) \\ 0 & (x < 0) \end{cases}$$

で表されているとき X はパラメータ λ の**指数分布**に従うという。

例 指数分布の期待値 $E(X)$ は次のように計算できます。

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} xf(x)dx = \int_{-\infty}^0 x \cdot 0dx + \int_0^{\infty} x\lambda e^{-\lambda x}dx = \int_0^{\infty} x(-e^{-\lambda x})'dx \\ &= [x(-e^{-\lambda x})]_{x=0}^{\infty} - \int_0^{\infty} (-e^{-\lambda x})dx = (0 - 0) - \left[\frac{1}{\lambda} e^{-\lambda x} \right]_{x=0}^{\infty} = \frac{1}{\lambda} \end{aligned}$$

また、分散はまず $E(X^2)$ を

$$\begin{aligned} E(X^2) &= \int_0^{\infty} x^2 f(x)dx = \int_0^{\infty} x^2 \lambda e^{-\lambda x}dx = \int_0^{\infty} x^2 (-e^{-\lambda x})'dx \\ &= [x^2(-e^{-\lambda x})]_{x=0}^{\infty} - \int_0^{\infty} 2x(-e^{-\lambda x})dx = 2 \int_0^{\infty} xe^{-\lambda x}dx = \frac{2}{\lambda^2} \end{aligned}$$

と求めて、公式 $V(X) = E(X^2) - (E(X))^2$ から

$$V(X) = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}$$

が得られます。

ガウスは誤差のモデルを考え、そのモデルを解析することで、誤差の従う分布の密度関数を $\phi(x)$ としたときに次の微分方程式

$$\phi'(x) = -x\phi(x)$$

を導きました。この解は次のように求めることができます。

$$\frac{\phi'(x)}{\phi(x)} = -x, \quad \log \phi(x) = -\frac{x^2}{2} + c, \quad \therefore \phi(x) = Ce^{-\frac{x^2}{2}} \quad (C = e^c)$$

これを $\int_{-\infty}^{\infty} f(x)dx = 1$ を満たすように定数 C を定めることにより次の分布の特別な場合が得られます.

定義 (正規分布).

μ を実数の定数, σ を正の定数とする. このとき連続型確率変数 X の密度関数が

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

で表されるとき X は平均 μ , 分散 σ^2 の**正規分布**に従うといい, $N(\mu, \sigma^2)$ で表す. 特に $N(0, 1)$ のことを**標準正規分布**という.

定理. X を正規分布 $N(\mu, \sigma^2)$ に従う確率変数とすると, $E(X) = \mu$, $V(X) = \sigma^2$.

定理. (1) 確率変数 X が $N(\mu, \sigma^2)$ に従うとき $aX + b$ は正規分布 $N(a\mu + b, a^2\sigma^2)$ に従う.
(2) 確率変数 X と Y が独立にそれぞれ $N(\mu_1, \sigma_1^2)$, $N(\mu_2, \sigma_2^2)$ に従うとき, $X + Y$ は正規分布 $N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$ に従う.

次の操作は基準化と呼ばれよく出てきます.

定理 (サンプル問題). 確率変数 X は母平均 μ , 母分散 σ^2 の正規分布に従うとする. このとき

$$Y = \frac{X - \boxed{}}{\boxed{}}$$

と変数変換すれば, 確率変数 Y の従う分布は母平均が 0, 母分散が 1 の正規分布に従う.

問 (2006 年統一試験). 確率変数 X が正規分布 $N(1, 2)$ に従っているとき, X の確率密度関数を $f(x)$ とすると, $\int_{-\infty}^{\infty} xf(x)dx = \boxed{}$. また X の分散 $V(X) = \boxed{}$ であるから, $\int_{-\infty}^{\infty} x^2 f(x)dx = \boxed{}$.

問 (2004 年統一試験). 確率変数 X が正規分布 $N(1, 4)$ に従っているとき, $Y = \frac{X-1}{2}$ とおくと, Y の期待値は $\boxed{}$ であり, Y の分散は $\boxed{}$ である.

問 (2005 年統一試験). 確率変数 X, Y は独立で共に平均 3 および分散 1 の正規分布 $N(3, 1^2)$ に従うものとする. このとき $X - Y$ の平均は $\boxed{}$ であり, 分散は $\boxed{}$ である.

2 推定と検定

2.1 統計量の分布

この節では後の推定・検定に用いる統計量の分布についてまとめる。最初は読み飛ばして先に推定・検定の事項を読む方がいいかもしれない。

定理. $X_1, \dots, X_n$ を独立に $N(\mu, \sigma^2)$ に従う確率変数とする。このとき $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ は $N(\mu, \frac{\sigma^2}{n})$ に従う。

物の長さを測定することを考えます。測定には誤差がつきものですが誤差の分布にある仮定をすると、測定値の分布は真の長さ μ を平均とした正規分布になります。このとき上の定理は n 回測定して平均をとったものの分散が、1 回だけの測定したものの分散の $1/n$ になることを示しています。つまり $\bar{X}$ のほうが μ に近い値が出やすいということです。

定義 (χ^2 分布) . —

$X_1, \dots, X_n$ を独立に $N(0, 1)$ に従う確率変数とする。このとき $Y = X_1^2 + \dots + X_n^2$ の従う分布を **自由度 n の χ^2 分布** という。

問 (サンプル問題). 確率変数 X^2 の従う分布が自由度 1 のカイ 2 乗分布に従うのは、 X の従う分布が母平均 0、母分散 1 の正規分布であるときに限る。【Yes, No】

定義 (t -分布) . —

X を $N(0, 1)$ に従う確率変数、 W を自由度 n の χ^2 分布に従う確率変数で、 X と W は独立であるとする。このとき $T = \frac{X}{\sqrt{W/n}}$ の従う分布を **自由度 n の t -分布** という。

t -分布の定義はこのとおりなのですが、統計学では次の形でよく出てきます。

定理. $X_1, \dots, X_n$ を独立に同一の正規分布 $N(\mu, \sigma^2)$ に従う確率変数とする。また $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$, $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ とおく。このとき

$$T = \frac{\bar{X} - \mu}{\sqrt{S^2/n}}$$

は自由度 $(n-1)$ の t -分布に従う。

定義 (F -分布).

V, W を独立にそれぞれ自由度 m, n の χ^2 分布に従う確率変数とする. このとき $F = \frac{V/m}{W/n}$ の従う分布を **自由度 m, n の F -分布** という.

問 (サンプル問題). t 分布に従う確率変数を 2 乗した確率変数は, 自由度が特別な場合の F 分布に従う. 【Yes, No】

2.2 点推定

点推定とは全体から一部を取り出して全体の平均や分散を一つの値で推定する方法である. 例を見てみよう.

例 内閣の支持率を調べるために国民全体から 10000 人をランダムに選んだところ 2000 人が支持すると答えた. 従って, 国民全体の内閣支持率も 20 % ぐらいであろう.

この節では, 選んだ標本からいかに推定をするかについて考える. その前に言葉の準備をしておく. 上の国民全体に相当するものを**母集団 (分布)** という. これには正規分布を仮定することが多い. 次に母集団からランダムに取り出された $X_1, \dots, X_n$ を**標本**という. そして $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ のように標本から計算される量 (確率変数) を**統計量**という. そして推定したい平均や分散のことを**母数**(パラメータ)という. 以下ではどのように母数を推定する統計量を作るかを考える.

まずモーメント法について簡単に触れる. 一般に $E(X), E(X^2), \dots, E(X^r), \dots$ のことを**モーメント**という. モーメント法は母数をモーメントで表現して, それに標本モーメント $\frac{1}{n} \sum_{i=1}^n X_i, \frac{1}{n} \sum_{i=1}^n X_i^2, \dots, \frac{1}{n} \sum_{i=1}^n X_i^r, \dots$ を代入して推定する方法である. 例を見てみよう.

例 母平均 μ は $\mu = E(X)$ と表されるので, $\hat{\mu} = \frac{1}{n} \sum_{i=1}^n X_i$ がモーメント法による μ の推定量である. また, 母分散 σ^2 は $E(X^2) - (E(X))^2$ と表されるので $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \left(\frac{1}{n} \sum_{i=1}^n X_i\right)^2$ がモーメント法による σ^2 の推定量である.

モーメント法は母数がモーメントで表現できない場合には使えない点に注意する. 次に, よく使われる最尤法について扱う. 例として二項分布 $B(n, p)$ の p の推定について考えよう.

例 画鋲を 10 回投げて「 $\perp$ 」の形かそれ以外の形かを調べる試行を行うと順に, 「他, 他, $\perp$, 他, 他, 他, $\perp$, $\perp$, 他, 他」となったとする. このとき「 $\perp$ 」となる確率を次のように推定しよう. 「 $\perp$ 」となる確率を p とすると, 上の標本が得られる確率は

$$L(p) = p^3(1-p)^{10-3}$$

で与えられる. これが最大となるような p を求めることを考える. そのまま $L(p)$ を p で微

分して $\frac{d}{dp}L(p) = 0$ となる p を求めてもよいが、計算の都合上 $\log L(p)$ を最大にする p を求める。(log は単調増加であるから $\log f(x)$ が $x = a$ で最大値をとるとき、 $f(x)$ も $x = a$ で最大値をとる.)

$$\frac{d}{dp} \log L(p) = \frac{d}{dp} (3 \log p + 7 \log(1-p)) = \frac{3}{p} - \frac{7}{1-p} = 0$$

これを解いて $\hat{p} = \frac{3}{10}$ を得る.

この例ように、「そのような標本が得られる確率(密度)」を最大にするような母数を推定量に使う方法を**最尤法**といい、その推定量を**最尤推定量**といいます.

一般の問題に使えるようにこれを次のように定式化しておきます.

定義 (最尤法). —

$X_1, X_2, \dots, X_n$ を確率(密度)関数が $f(x|\theta)$ であるような母集団からの独立標本であるとする. ここで θ は推定したいパラメータである. このとき

$$L(\theta) = \prod_{i=1}^n f(X_i|\theta) = f(X_1|\theta) \times \cdots \times f(X_n|\theta)$$

を**尤度関数**といい、これを最大にする $\theta = \hat{\theta}$ を**最尤推定量**という. なお、尤度関数の対数 $l(\theta) = \log L(\theta)$ の最大値を与える θ も $\hat{\theta}$ であるので

$$\frac{d}{d\theta} \log L(\theta) = 0$$

を θ について解けば最尤推定量が得られる.

問 (2005 年統一試験). 確率変数 X の密度関数が

$$f(x; \lambda) = \begin{cases} \frac{1}{\lambda} e^{-\frac{x}{\lambda}} & (x \geq 0 \text{ のとき}) \\ 0 & (x < 0 \text{ のとき}) \end{cases}$$

で与えられているとする. ただし、 $\lambda > 0$ は未知のパラメータである. X の n 個の観測値が $x_i \geq 0$ ($i = 1, 2, \dots, n$) であるときの λ の最尤推定量 $\hat{\lambda}$ を求めたい. 観測値の組を $\mathbf{x} = (x_1, \dots, x_n)$ と書くことにすると、 $\hat{\lambda}$ は尤度関数

$$L(\lambda; \mathbf{x}) = \left(\frac{1}{\lambda} e^{-\frac{x_1}{\lambda}} \right) \left(\frac{1}{\lambda} e^{-\frac{x_2}{\lambda}} \right) \cdots \left(\frac{1}{\lambda} e^{-\frac{x_n}{\lambda}} \right)$$

を最大にする λ として求められる. 計算を簡単にするために $L(\lambda; \mathbf{x})$ の自然対数を $l(\lambda; \mathbf{x})$ と

すると,

$$l(\lambda; \mathbf{x}) = -n \log \lambda - \left(\frac{\sum_{i=1}^n x_i}{\lambda} \right)$$

となり, 極値を求めるために λ で微分すると

$$l(\lambda; \mathbf{x}) = -\frac{n}{\lambda} + \left(\frac{\sum_{i=1}^n x_i}{\lambda} \right)$$

となる. ここで $\frac{dl}{d\lambda}(\lambda; \mathbf{x}) = 0$ を満たす λ を求め $l(\lambda; \mathbf{x})$ の増減を調べることにより, $\lambda =$

$\frac{\sum_{i=1}^n x_i}{n}$ で $l(\lambda; \mathbf{x})$ が最大になることが分かる. したがって最尤推定量 $\hat{\lambda} = \frac{\sum_{i=1}^n x_i}{n}$ である.

2.3 区間推定

内閣支持率の例でいうと, 点推定では 100 人中 10 人でも 1000 人中 100 人であっても同じ推定値 10 % を得ます. しかし, 直感的には後者の 1000 人中 100 人のほうが 10 % という値に信頼性があると感じられます. また, 国民全体の支持率が丁度 10 % に等しいかどうかは怪しいところがあります. そこで母数が含まれる区間を推定することを考えます. 次の例を見てください.

例 ある日のニュースの世論調査で最初に「予測最大誤差 2.6 %」とあった. このときの民主党支持の割合は 25.0 % だったので, 国民全体の民主党支持の割合は「22.4 ~ 27.6 %」に含まれているであろう.

このように母数が含まれる区間「 $a \leq \theta \leq b$ 」をデータから推定することを考えます.

例 あるネジの長さを 4 回測定したところ, 9.9, 10.0, 10.2, 10.3 であった. 過去の経験からネジの長さは平均 μ で分散が $\sigma^2 = 0.04$ の正規分布に従っていることが分かっているとする. このときネジの真の長さの信頼区間を次のように求めよう.

X_1, X_2, X_3, X_4 を正規分布 $N(\mu, 0.04)$ からの標本とすると, この標本平均 $\bar{X} = \frac{X_1 + X_2 + X_3 + X_4}{4}$ は正規分布 $N(\mu, \frac{0.04}{4})$ に従う. 従って $\bar{X}$ を基準化した

$$Z = \frac{\bar{X} - \mu}{\sqrt{0.04/4}}$$

を考えると Z は標準正規分布 $N(0, 1)$ に従う. これを利用して $P(a \leq \mu \leq b) = 0.95$ となるような区間 $[a, b]$ を求めよう. 数表によると $P(Z \geq 1.96) = 0.025$ で, 標準正規分布は y 軸を中心に左右対称であるから

$$P(-1.96 \leq Z \leq 1.96) = 0.95$$

が成り立っている．これに $Z = \frac{\bar{X} - \mu}{\sqrt{0.04/4}}$ を代入して μ について整理すると

$$P\left(\bar{X} - 1.96\sqrt{\frac{0.04}{4}} \leq \mu \leq \bar{X} + 1.96\sqrt{\frac{0.04}{4}}\right) = 0.95$$

が得られる．従って実際の標本に対しても $P(\cdot)$ の不等式が成り立っていると期待できる．この不等式に $\bar{x} = (9.9 + 10.0 + 10.2 + 10.3)/4 = 10.1$ を代入した $[10.0, 10.2]$ のことを μ の **95%信頼区間** という．

定義&公式 (母平均の信頼区間) .

(I) (母分散 σ^2 が既知のときの μ の信頼区間). $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ とする．このとき

$$Z = \frac{\bar{X} - \mu}{\sqrt{\sigma^2/n}}$$

は標準正規分布に従う． $P(-z_{\alpha/2} \leq Z \leq z_{\alpha/2}) = 1 - \alpha$ の $P(\cdot)$ を変形して次の $100 \times (1 - \alpha)$ %信頼区間が得られる．

$$\left[\bar{X} - z_{\alpha/2} \sqrt{\frac{\sigma^2}{n}}, \quad \bar{X} + z_{\alpha/2} \sqrt{\frac{\sigma^2}{n}} \right]$$

ここで $z_{\alpha/2}$ は $Z \sim N(0, 1)$ としたときの $P(Z \geq z_{\alpha/2}) = \alpha/2$ の値を表している．よく使われる値を表にしておく．

$P(Z \geq z_{\alpha})$	0.005	0.01	0.025	0.05	0.10
z_{α}	2.576	2.326	1.960	1.645	1.282

(II) (母分散 σ^2 が未知のときの μ の信頼区間). $V = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ とする．このとき

$$T = \frac{\bar{X} - \mu}{\sqrt{V/n}}$$

は自由度 $n - 1$ の t -分布に従うことが知られている．同様に信頼区間は

$$\left[\bar{X} - t_{n-1, \alpha/2} \sqrt{\frac{V}{n}}, \quad \bar{X} + t_{n-1, \alpha/2} \sqrt{\frac{V}{n}} \right]$$

で得られる．ここで $t_{n-1, \alpha/2}$ は T_{n-1} を自由度 $n - 1$ の t -分布に従う確率変数としたときの $P(T_{n-1} \geq t_{n-1, \alpha/2}) = \alpha/2$ となる値でこれも数表から求めることができる．

問 (2005 年統一試験). 正規分布の母平均を区間推定するとき、信頼度 (信頼係数) を変えずに標本の大きさを大きくすると信頼区間は【狭くなる・広くなる・変らない】．また、標本の大きさを変えずに信頼度を大きくすると信頼区間は【狭くなる・広くなる・変らない】．

参考 (ド・モアブル＝ラプラスの定理). X を二項分布 $B(n, p)$ に従う確率変数とするとき次が成り立ちます.

$$P\left(\frac{X - np}{\sqrt{np(1-p)}} \leq x\right) \rightarrow \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du \quad (n \rightarrow \infty)$$

これは n が大きいときの二項分布が正規分布で近似できることを示しています. これを用いて冒頭の支持率 p の信頼区間は計算されているものと思われます.

問 (2006 年統一試験). A 大学のキャンパスの池にはウシガエルが多数生息している. U 教授の研究室では毎年この池のウシガエルの調査を行っている. 今までの調査によって, この池に生息するウシガエルの体長の分布は正規分布に従い, 体長の母分散 σ^2 は年によって変化はなく, $\sigma^2 = 35^2 \text{mm}^2$ であることがわかっている. 今年も 100 匹捕獲して体長を測定したところ, 100 匹の標本平均値は $\bar{x} = 165 \text{mm}$ であった. そこでこの池に生息する全てのウシガエルの体長の母平均 μ の信頼度 (信頼係数) 95 % の信頼区間を以下のように求めた.

捕獲した 100 匹のウシガエルの体長を表す確率変数をそれぞれ $X_1, X_2, \dots, X_{100}$ とおくと, これらは独立で, 全て平均 μ , 分散 35^2 の正規分布 $N(\mu, 35^2)$ に従っている. ゆえに標本平均

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_{100}}{100}$$

の分布は平均 , 分散 $\frac{\sigma^2}{100} = \frac{35^2}{100}$ の正規分布である. そこで

$$Z = \frac{\bar{X} - \text{}}{\text{}}$$

とおけば, Z の分布は標準正規分布 $N(0, 1)$ である. ここで正規分布表を調べると, およそ

$$P(-1.96 \leq Z \leq 1.96) = 0.95 \quad (*)$$

であることがわかった. 式 (*) を書きかえると

$$P\left(\bar{X} - 1.96 \times \frac{\text{}}{\text{}} \leq \mu \leq \bar{X} + 1.96 \times \frac{\text{}}{\text{}}\right) = 0.95$$

となる. これから実際の標本平均値 $\bar{x}$ に対しても不等式

$$\bar{x} - 1.96 \times \frac{\text{}}{\text{}} \leq \mu \leq \bar{x} + 1.96 \times \frac{\text{}}{\text{}}$$

が成り立っていると推定する. これに数値を代入して計算した結果の小数点以下第 1 位を四捨五入すると, 求める母平均 μ の信頼度 95 % の信頼区間は (mm) $\leq \mu \leq$ (mm) である.

2.4 仮説検定

仮説検定はデータからパラメータに関する主張を判断することを行います。例を見て見ましょう。

例 S社の競泳用水着を着用した水泳選手達が世界新・日本新記録を含む自己ベストを連発している。S社の水着が優れているのかを調べるためにランダムに選ばれた10人にS社と他社の水着をそれぞれ着用させ、100m自由形の記録をとった。このデータからS社の水着を着用したほうが記録が向上することが言えるであろうか。

このような問題を考えていきます。

例 ある日のニュースによると今の高校生の足の長さ（股下から足首までの長さ）は30年前の高校生の足の長さよりも短いそうである。これを検証してみよう。30年前の高校生の足の長さの平均を $\mu_0(\text{cm})$ として、今の高校生の足の長さの平均を $\mu(\text{cm})$ とする。30年前の足の長さの平均 μ_0 は分かっているとして、今の高校生をランダムに n 人選び、足の長さ $X_1, \dots, X_n$ とする。また、簡単のため足の長さは正規分布 $N(\mu, \sigma^2)$ に従っていて、 σ^2 は既知とする。このとき考えたい問題はデータから

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

のどちらを判断するかである。 H_0 は「30年前と変っていない」、 H_1 は「30年前よりも短くなった」の意味である。ここで $\bar{X}$ の分布は $N(\mu, \sigma^2/n)$ であることから次が考えられる。

- H_0 が正しいとすると $\bar{X}$ は μ_0 に近い値をとりやすい。
- H_1 が正しいとすると $\bar{X}$ は μ_0 よりも小さな値をとりやすい。

そこである閾値 $-c < 0$ を用いて、実際に得られた標本平均 $\bar{x}$ に対して

- $\bar{x} \geq \mu_0 - c$ ならば H_0 と判断する。(これを H_0 を**採択する**という.)
- $\bar{x} < \mu_0 - c$ ならば H_1 と判断する。(これを H_0 を**棄却する**という.)

という方法が考えられる。ではこの閾値 c をどのように定めればよいであろうか。このとき注意することは c をいくつにしても、 H_0 が正しいときの $P(\bar{X} < \mu_0 - c)$ は0にならないことである。つまり、30年前と比べて変ってないとしても、たまたま $\bar{X}$ が小さい値をとることがあるということである。そこで

$$P(\bar{X} < \mu_0 - c) = 0.05 \quad (H_0 \text{が正しいとき})$$

を満たす $c = c_{0.05}$ を用いることを考える。つまり、 H_0 が正しいときに間違った判断を下すことも許すが、それが起きる確率を0.05という小さな値に押さえることを考える。ここで H_0 が正しいとき $\bar{X}$ は $N(\mu_0, \sigma^2)$ に従うので、基準化した

$$Z = \frac{\bar{X} - \mu_0}{\sqrt{\sigma^2/n}}$$

は標準正規分布 $N(0, 1)$ に従う. 数表によると $P(Z < -1.645) = 0.05$ なのでこれに $Z = \frac{\bar{X} - \mu_0}{\sqrt{\sigma^2/n}}$ を代入して整理すると

$$P\left(\bar{X} < \mu_0 - 1.645\sqrt{\frac{\sigma^2}{n}}\right) = 0.05$$

となる. よって $\bar{x}$ が $\mu_0 - 1.645\sqrt{\frac{\sigma^2}{n}}$ が大きいときは H_0 を採択し, 小さいときは H_0 を棄却する.

注意. 仮説検定は内心, 帰無仮説 H_0 が棄却されることを望んでいます. しかし, 棄却されなかった場合は必ずしも帰無仮説が正しいとは限りません. それは上の論法では「帰無仮説が正しい」ときの確率を元に判断を下しているので「対立仮説が正しい」ときの情報は何も使っていないからです. 詳しくは統計学の参考書で「検出力」を調べてみてください.

検定で現れるいくつかのパターンをまとめておきます.

定義&公式 (母分散 σ^2 が既知のときの母平均に関する仮説検定).

(I-1) (両側検定). 次の検定問題

$$H_0 : \mu = \mu_0 \quad \text{v.s.} \quad H_1 : \mu \neq \mu_0$$

を考える. 次の統計量

$$Z = \frac{\bar{X} - \mu_0}{\sqrt{\sigma^2/n}}$$

は帰無仮説 H_0 の元で標準正規分布 $N(0, 1)$ に従うので次の検定方式が得られる.

$$|Z| > z_{\alpha/2} \text{ のとき } H_0 \text{ を棄却, } |Z| \leq z_{\alpha/2} \text{ のとき } H_0 \text{ を採択}$$

(I-2) (片側検定). 次の検定問題

$$H_0 : \mu = \mu_0 \quad \text{v.s.} \quad H_1 : \mu > \mu_0$$

$$(H_0 : \mu = \mu_0 \quad \text{v.s.} \quad H_1 : \mu < \mu_0)$$

を考える. 次の統計量

$$Z = \frac{\bar{X} - \mu_0}{\sqrt{\sigma^2/n}}$$

は帰無仮説 H_0 の元で標準正規分布 $N(0, 1)$ に従うので次の検定方式が得られる.

$$\begin{aligned} Z > z_{\alpha} \text{ のとき } H_0 \text{ を棄却, } Z \leq z_{\alpha} \text{ のとき } H_0 \text{ を採択} \\ (Z < -z_{\alpha} \text{ のとき } H_0 \text{ を棄却, } Z \geq -z_{\alpha} \text{ のとき } H_0 \text{ を採択}) \end{aligned}$$

定義&公式 (母分散 σ^2 が未知のときの母平均に関する仮説検定).

(II-1) (**両側検定**). 次の検定問題

$$H_0 : \mu = \mu_0 \quad \text{v.s.} \quad H_1 : \mu \neq \mu_0$$

を考える. 次の統計量

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

は帰無仮説 H_0 の元で自由度 $n - 1$ の t -分布に従うので次の検定方式が得られる.

$$|T| > t_{n-1, \alpha/2} \text{ のとき } H_0 \text{ を棄却, } |T| \leq t_{n-1, \alpha/2} \text{ のとき } H_0 \text{ を採択}$$

(II-2) (**片側検定**). 次の検定問題

$$\begin{aligned} H_0 : \mu = \mu_0 \quad \text{v.s.} \quad H_1 : \mu > \mu_0 \\ (H_0 : \mu = \mu_0 \quad \text{v.s.} \quad H_1 : \mu < \mu_0) \end{aligned}$$

を考える. 次の統計量

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

は帰無仮説 H_0 の元で自由度 $n - 1$ の t -分布に従うので次の検定方式が得られる.

$$\begin{aligned} T > t_{n-1, \alpha} \text{ のとき } H_0 \text{ を棄却, } T \leq t_{n-1, \alpha} \text{ のとき } H_0 \text{ を採択} \\ (T < -t_{n-1, \alpha} \text{ のとき } H_0 \text{ を棄却, } T \geq -t_{n-1, \alpha} \text{ のとき } H_0 \text{ を採択}) \end{aligned}$$

問 (2004 年統一試験). 仮説検定を行って帰無仮説 H_0 が棄却されなかったならば H_0 は正しい. 【Yes, No】

問 (サンプル問題). 帰無仮説が正しいときにこの帰無仮説を棄却する確率と, 帰無仮説が正しくないときにこの帰無仮説を棄却しない確率はどちらも同じである. 【Yes, No】

問 (サンプル問題). 仮説検定を行うときは, 事前に有意水準を決めておくことが望ましい. 【Yes, No】

問 (2007 年統一試験). A 大学の U 教授の講義は面白くて人気がある. 以前は, 受講者の $\frac{4}{5}$ は単位を取れると学生の間で語り継がれてきた. しかし最近は単位認定が厳しくなったと噂されている. 平成 19 年度の単位取得者は受講者 100 人のうち 73 人であった. U 教授の単位認定は以前よりも厳しくなったといえるだろうか? これを判断するため, 帰無仮説を「U 教授は受講者の $\frac{4}{5}$ の単位を認定する」として有意水準 (危険率) 5% で左片側検定を実施する. まず, 受講者 100 人のうち番号 n ($n = 1, 2, \dots, 100$) の学生について, 単位を取れば

$X_n = 1$, 取れなければ $X_n = 0$ とおく. このとき, 帰無仮説の下で統計量 $S = \sum_{n=1}^{100} X_n$ は 2 項分布に従い $E(S) = \square$, $V(S) = \square$ である. 次に, 平均が 0, 分散が 1 になるように S を標準化した統計量を $T = \frac{S - \square}{\square}$ とすると, 平成 19 年度の単位認定に対する T の実現値 t は $\square$ となる. T の分布は平均 0, 分散 1 の正規分布 (標準正規分布) で近時されるので, $P(T \geq 1.654)$ はおよそ 0.05 であることがわかる. したがって, $\square$ から, 左片側検定において帰無仮説は有意水準 (危険率) 5 % で【棄却される, 棄却されない】.

参考. 冒頭の水着の問題はもう少し複雑です. それは「この水着は新作で高記録が出るに違いない」という思い込みで高記録が出ることがあるからです. これは新薬の効果を確かめるときなどにも現れる問題で心理的な要因を排除する様々な方法が考えられています.

3 補遺

概要にも述べたが本稿は確率統計の本としては不十分である. この補遺では少しでも内容の補充を図る.

定義 (積率母関数). —

確率変数 X に対して次の期待値を考える.

$$M(t) = E(e^{tX}) = \begin{cases} \sum e^{tx} p(x) \\ \int_{-\infty}^{\infty} e^{tx} f(x) dx \end{cases}$$

これを X の**積率母関数**という.

なぜこのようなものを考えるかということですが, 一つの理由としてモーメントが簡単に計算できることがあります. 上の連続型の場合の e^{tX} をテイラー展開すると (ある条件下で)

$$\begin{aligned} M(t) &= E \left[1 + tX + \frac{1}{2}t^2 X^2 + \cdots + \frac{1}{n!}t^n X^n + \cdots \right] \\ &= 1 + tE(X) + \frac{t^2}{2}E(X^2) + \cdots + \frac{t^n}{n!}E(X^n) + \cdots \end{aligned}$$

が成り立ちます. $M(t)$ を t で 1 回微分すると

$$M'(t) = E(X) + tE(X^2) + \cdots + \frac{t^{n-1}}{(n-1)!}E(X^n) + \cdots$$

となりこれに $t = 0$ を代入すると $M'(0) = E(X)$ となり、 X の期待値が得られます。また 2 回微分して $t = 0$ を代入すると $E(X^2) = M''(0)$ が得られ、これから分散を

$$V(X) = M''(0) - (M'(0))^2$$

と求めることができます。

前にも二項分布が正規分布で近似できることに触れたが実はもっと強いことが言える。

定理 (中心極限定理). $X_1, \dots, X_n$ を独立に母平均 μ , 母分散 σ^2 である母集団に従う確率変数とする. このとき $Z = \frac{\bar{X} - \mu}{\sqrt{\sigma^2/n}}$ と基準化すると, Z の分布関数 $F(z)$ について $n \rightarrow \infty$ のとき

$$F(z) = P(Z \leq z) \rightarrow \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

となる.

この定理の凄いところは例えば母集団が正規分布でなくても、標本数 n が大きいときは $\bar{X}$ の分布が $N(\mu, \sigma^2/n)$ で近似できるというところにある。

次に §2.2 で推定量の作り方を紹介したが、推定量が満たしたほうがいい望ましい性質を紹介する。

定義 (不偏推定量). —

母数が μ である母集団から抽出された独立な標本 $X_1, X_2, \dots, X_n$ を使って推定量 $\hat{\mu} = \hat{\mu}(X_1, \dots, X_n)$ を作ったとする. もし $\hat{\mu}$ が $E(\hat{\mu}) = \mu$ を満たすとき, $\hat{\mu}$ を**不偏推定量**という.

例 標本平均 $\bar{X}$ について

$$E(\bar{X}) = E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} \sum_{i=1}^n E(X_i) = \frac{1}{n} \sum_{i=1}^n \mu = \frac{1}{n} n\mu = \mu$$

であるから, $\bar{X}$ は μ の不偏推定量である.

不偏推定量は何回も推定量を作ったときに, それらの平均が推定したい母数に収束することを言っている.

他にも大数の法則, 確率変数の変数変換や推定・検定の問題補充などをしたいところではあるが, あまり膨大になっても読みにくいと思うのでこの辺りにしておく. 教科書・参考書で内容の補充や問題演習などを積んでほしい.