

IT 資格 道 場

統計検定準1級 学習用テキスト

予備知識ゼロから合格ラインへ —— 読んで理解する1冊目

無料サンプル(第1章まで)

統計検定 公式「統計検定準1級 出題範囲表」(2027年1月1日改訂版・公式ページ表記
202107版) 準拠
2026年10月版

ご利用にあたって

- このテキストは「IT資格道場」(www.shikaku-doujou.com)の購入者限定テキストのうち、学習用テキストの第1章までを収録した**無料サンプル**です。どなたでもご覧いただけます。
- 無料サンプルも著作物です。転載・改変しての再配布・二次販売はできません。
- AI モデルの学習データとしてのご利用(学習目的の入力・データセット化を含む)はできません。
- 製品版(学習用テキスト・頻出度別ノート・暗記用ノートの3点)は、ご購入者を識別できる透かし入りのPDFとして提供されます。
- 本PDFはスマートフォン・タブレットでの閲覧に合わせたA5版面・文字サイズで組版しています。

目次

読み始める前に:この教材の使い方

試験の基本情報

4 ステップ学習法

第1章 確率と確率変数 [1.1/1.2/1.3/1.4/1.5]

事象と確率 [1.1]

確率分布と母関数 [1.2]

分布の特性値 [1.3]

変数変換 [1.4]

極限定理, 漸近理論 [1.5]

サンプルはここまでです

3冊の使い方は、ダウンロード欄の案内を参照してください。

SAMPLE

読み始める前に:この教材の使い方

このテキストは、一般財団法人 統計質保証推進協会が実施する **統計検定準1級** に合格することを目的に作られています。

準1級が問うのは、2級の内容をすべて土台にして、確率と確率変数の理論、各種の統計解析法の使い方、そして解析結果の正しい解釈です。

範囲表は 2027年1月1日から新しい版に替わり、追加される大項目は「統計的因果推論」だけです(2026年12月31日までの試験は旧範囲表で実施)。第1章～第16章は、この範囲表の大項目の順に並べました。

本番では、計算を要する問題を避けて通れません。各節の計算の型は、式、数値例、答えの順。式が出てきたら、手を動かして数値例をなぞってください。

試験の基本情報

項目	内容
試験名	統計検定準1級
運営	一般財団法人 統計質保証推進協会(日本統計学会公式認定)
受験方法	CBT 方式(全国のおデッセイの試験会場)
出題形式	5肢選択問題、数値入力問題
問題数	25問～30問 (公式ページ)
試験時間	90分
合格水準	100点満点で、60点以上

項目	内容
持ち込み	四則演算・百分率・平方根が計算できる普通電卓か事務用電卓を1台まで(関数電卓・プログラム電卓などは不可)
配布物	統計数値表の付表1～4、計算用紙、筆記用具を会場で配り、試験の後に回収

※ 分野別の配点は公式には公表されていません。

4 ステップ学習法

1. 第1章から順に読む
2. 各節の数値例を、電卓を使って手を動かして確かめる
3. 問題演習で腕試しをし、間違えた問題の該当節に戻る
4. 頻出度別ノートで直前の総仕上げ

第1章 確率と確率変数 [1.1/1.2/1.3/1.4/1.5]

準1級の出題範囲表は、大項目「確率と確率変数」に5つの小項目を置いています。[^c1-1] 小項目は、事象と確率、確率分布と母関数、分布の特性値、変数変換、極限定理・漸近理論の順です。2級の内容はすべて含まれるため、加法定理、期待値と分散の基本、正規分布の標準化は前提です。本文で一から教え直すことはありません。

この章の計算は、四則演算・百分率・平方根だけの普通電卓で解けます。指数関数や対数が要る計算は普通電卓では解けないため、本文では値を与えた形で読みます(例: $e^{-1.5} = 0.2231$ 、 $\ln 2 = 0.6931$)。正規分布の確率は、配布される数値表で引きます。本文では上側確率 $P(Z > z)$ の形で書きます。[^c1-2] 本文の確率は、 z を小数第2位に丸めて付表で引いた値です(例: $z = 1.20$ の上側確率 0.1151)。

事象と確率 [1.1]

この節の語は、確率の計算、統計的独立、条件付き確率、ベイズの定理、包除原理の5つです。出題は、独立と排反の取り違え、条件の向きの取り違え、重なりの数え漏れをねらいます。式を暗記するより、問題文の言葉を式の記号へ直す手順を固めない限り、点は伸びません。

1.1-1 確率の基本性質と計算

確率は、標本空間 Ω の部分集合である事象 A に、 0 以上 1 以下の数を割り当てる関数です。 $P(\Omega) = 1$ と、排反な事象の和の確率が確率の和になる性質が土台です。加法定理と余事象の式は次のとおりです。

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$P(A^c) = 1 - P(A)$$

場合の数を全体の根元事象数で割る計算は、根元事象が同様に確からしいときだけ使えます。サイコロを2個振って目の和が7になる確率は、36通り中6通りで $1/6$ です。目の和が 2 から 12 の11通りを等確率とみなすことは、できません。和は根元事象ではなく、根元事象をまとめた事象だからです。

「少なくとも1つ」は余事象で数えます。硬貨を5回投げて少なくとも1回表が出る確率は、余事象から次のように出ます。

$$1 - (1/2)^5 = 31/32$$

「少なくとも」を見たら、「1つも起きない確率」を先に出してください。

間違えやすいのは、 $P(A \cup B)$ を $P(A) + P(B)$ で終えてしまうことです。重なり $P(A \cap B)$ を引き忘れると、確率が1を超えることもあるので要注意。

1.1-2 統計的独立と排反の区別・独立性の判定

事象 A と B が統計的に独立であるとは、次の式が成り立つことです。排反であるとは、 $A \cap B$ が空集合で $P(A \cap B) = 0$ となることです。

$$\text{独立: } P(A \cap B) = P(A)P(B)$$

この2つは別の性質です。 $P(A) > 0$ かつ $P(B) > 0$ なら、排反な2事象は独立になれません。 $P(A \cap B) = 0$ に対して $P(A)P(B) > 0$ だからです。独立は「片方が起きても、もう片方の起きやすさが変わらない」という意味で、排反はむしろ強い従属であり、独立の仲間ではありません。

3つ以上の事象では、相互独立と対独立を分けます。相互独立では、どの部分集合でも積の形が成り立ちます。A, B, C なら、3つの対の積と $P(A \cap B \cap C) = P(A)P(B)P(C)$ の4本がそろう必要があります。対ごとの独立(対独立)だけでは足りません。

反例は硬貨2枚です。A を「1枚目が表」、B を「2枚目が表」、C を「2枚が同じ面」とすると、 $P(A) = P(B) = P(C) = 1/2$ で、どの2つの積も $1/4$ に一致します。対独立です。ところが $A \cap B \cap C$ は「2枚とも表」で確率 $1/4$ であり、積の $1/8$ とは違います。相互独立ではありません。

独立性の判定は、表から $P(A \cap B)$ を読み、 $P(A)P(B)$ と比べて行います。一致しなければ従属です。独立なら、 A と B^c 、 A^c と B^c の組も独立です(補事象に替えても崩れません)。

数値例です。 $P(A) = 0.4$ 、 $P(B) = 0.5$ 、 $P(A \cup B) = 0.7$ のとき、加法定理から次のように判定します。

$$P(A \cap B) = 0.4 + 0.5 - 0.7 = 0.2 = 0.4 \times 0.5$$

積に一致するので、独立です。 $P(A \cup B) = 0.8$ なら $P(A \cap B) = 0.1$ で、積の 0.2 とは違うため、従属になります。

1.1-3 条件付き確率と乗法定理

事象 B が起きたという条件のもとで A が起きる確率を、条件付き確率といいます。定義と乗法定理は次の形です。

$$P(A|B) = P(A \cap B) / P(B) \quad (P(B) > 0)$$

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

A と B が独立なら $P(A|B) = P(A)$ です。条件を知っても確率が変わらない、というのが独立の言い換えです。

間違えやすいのは向きです。 $P(A|B)$ と $P(B|A)$ は一般に等しくありません。「陽性の人が病気である確率」と「病気の人が陽性になる確率」は別の量で、前者は陽性の側、後者は病気の側で割ります。

袋に赤玉3個と白玉2個があり、戻さずに2個取り出すとします。1個目が赤で2個目も赤の確率は、乗法定理から次のとおりです。

$$(3/5) \times (2/4) = 3/10$$

2個目が赤であるという条件のもとで1個目も赤である確率は、2個目が赤の確率 $3/5$ で割って $(3/10) / (3/5) = 1/2$ です。条件の向きを変えると、分母が変わります。

1.1-4 ベイズの定理(事後確率・検査の陽性的中率)

$A_1, \dots, A_k$ が標本空間の分割(互いに排反で、和集合が全体)であるとき、全確率の公式とベイズの定理が成り立ちます。

$$P(B) = \sum_j P(B|A_j)P(A_j)$$

$$P(A_i|B) = P(B|A_i)P(A_i) / \sum_j P(B|A_j)P(A_j)$$

$P(A_i)$ を事前確率、 $P(A_i|B)$ を観測 B を得たあとの事後確率といいます。観測から原因の確からしさを更新する計算です。

典型は検査の陽性的中率です。有病率 1%、感度(病気の人が陽性になる確率)90%、偽陽性率(健康な人が陽性になる確率)5% の検査で、陽性者のうち本当に病気の人割合は次のとおりです。

$$P(\text{病気}|\text{陽性}) = 0.01 \times 0.90 / (0.01 \times 0.90 + 0.99 \times 0.05) = 0.009 / 0.0585 \\ = 0.1538$$

陽性なのに病気である確率は、約15%にとどまります。分母の大部分を、健康な人の偽陽性(0.0495)が占めるからです。有病率が低いほど、陽性的中率は感度よりずっと小さくなります。

手順は3つです。

- ①分割(病気・健康)と観測(陽性)を決める。
- ②事前確率と、各原因のもとで観測が起きる確率を表に書く。
- ③分子に注目する原因の積、分母に全原因の積の和を置く。

②の表を先に作ると、向きの取り違えが消えます。

分割が3つ以上の場面も同じ手順です。工場 A, B, C が全体の 50%, 30%, 20% を作り、不良率が 2%, 3%, 5% とします。不良品を1個選んだとき、それが C の製品である確率は次のとおりです。

$$P(\text{不良}) = 0.5 \times 0.02 + 0.3 \times 0.03 + 0.2 \times 0.05 = 0.029 \\ P(C|\text{不良}) = 0.01/0.029 = 0.345$$

検査を重ねる場面では、事後確率が次の事前確率になります。上の検査で2回続けて陽性(結果は病気の有無のもとで独立)なら、1回目の事後確率 0.1538 を事前確率として使います。

$$0.1538 \times 0.90 / (0.1538 \times 0.90 + 0.8462 \times 0.05) = 0.1385/0.1808 = 0.766$$

2回陽性で、約77%まで上がります。

1.1-5 包除原理(3つ以上の事象の和の確率)

2事象の加法定理を3事象へ広げた式が、包除原理です。

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$

足して、2つ重なった部分を引き、3つ重なった部分を足し戻す、という符号の交代が規則です。n 事象では、k 個の積事象の確率の和に $(-1)^{k+1}$ を掛け、k = 1 から n まで足します。

数値例です。P(A) = 0.5、P(B) = 0.4、P(C) = 0.3、P(A ∩ B) = 0.2、P(A ∩ C) = 0.15、P(B ∩ C) = 0.12、P(A ∩ B ∩ C) = 0.08 のとき、次の値になります。

$$P(A \cup B \cup C) = 1.2 - 0.47 + 0.08 = 0.81$$

最後の +P(A ∩ B ∩ C) を忘れると 0.73 になります。3つ重なった部分は、足すときに3回、引くときに3回数えられて0回になるので、1回足し戻さなければ合いません。使い所は「少なくとも1つ」の確率です。1 から引けば、3事象のどれも起きない確率が出ます。

場合の数にも使えます。1 から 100 までの整数のうち、2, 3, 5 の少なくとも1つで割り切れる数の個数は、次のとおりです。

$$50 + 33 + 20 - 16 - 10 - 6 + 3 = 74$$

2 と 3 の公倍数は 6 の倍数で 16 個、2 と 5 は 10 の倍数で 10 個、3 と 5 は 15 の倍数で 6 個、3つすべては 30 の倍数で 3 個です。

確率分布と母関数 [1.2]

この節の語は、次の12個です。

- 確率関数、確率密度関数
- 同時確率関数、同時確率密度関数
- 周辺確率関数、周辺確率密度関数
- 条件つき確率関数、条件つき確率密度関数
- 累積分布関数、生存関数
- モーメント母関数(積率母関数)、確率母関数

同時分布から周辺分布と条件つき分布を取り出す操作が、この節の中心です。

1.2-1 確率関数・確率密度関数の条件(非負・全確率1・規格化定数)

離散確率変数 X の確率関数 $f(x) = P(X = x)$ は、 $f(x) \geq 0$ と $\sum f(x) = 1$ を満たします。期待値は次の式です。

$$E[X] = \sum x f(x), \quad E[g(X)] = \sum g(x) f(x)$$

連続確率変数 X の確率密度関数は、 $f(x) \geq 0$ と $\int f(x) dx = 1$ を満たします。確率は面積で与えられます。

$$P(a \leq X \leq b) = \int_a^b f(x) dx, \quad P(X = x) = 0, \quad F'(x) = f(x)$$

密度 $f(x)$ の値そのものは確率ではないので、1を超えてかまいません。区間 $(0, 0.5)$ 上の一様分布の密度は2です。連続変数では1点の確率が0なので、不等号に等号が付いても結果は変わりません。

規格化定数は、全確率が1になる条件から決めます。 $f(x) = c x^2$ ($0 \leq x \leq 3$) なら、 $\int_0^3 c x^2 dx = 9c = 1$ より $c = 1/9$ です。続けて、確率と期待値が出ます。

$$P(1 < X < 2) = (1/9) \int_1^2 x^2 dx = (8 - 1)/27 = 7/27$$

$$E[X] = (1/9) \int_0^3 x^3 dx = 9/4$$

間違えやすいのは、台の外で f を 0 と置く約束です。積分は、密度が正の値をとる範囲の中だけで行い、範囲の外へはみ出してはいけません。

1.2-2 累積分布関数と生存関数(密度との関係・ハザード関数)

累積分布関数は $F(x) = P(X \leq x)$ です。非減少で右連続です。性質と区間の確率は次の形です。

$$F(-\infty) = 0, \quad F(\infty) = 1, \quad P(a < X \leq b) = F(b) - F(a)$$

連続なら $f(x) = F'(x)$ です。離散なら F は階段関数で、ジャンプの大きさが $P(X = x)$ に当たります。

寿命 $T \geq 0$ には、生存関数とハザード関数を使います。

$$S(t) = P(T > t) = 1 - F(t)$$

$$h(t) = f(t) / S(t)$$

$$E[T] = \int_0^\infty S(t) dt$$

S は非増加で、 $S(0) = 1$ から 0 に向かいます。ハザード関数は、時刻 t まで生き延びた対象が、直後の短い時間に事象を起こす割合です。指数分布では $h(t) = \lambda$ で一定になります($S(t) = e^{(-\lambda t)}$ のとき)。経年劣化も初期不良もない状態を表します。ワイブル分布 $S(t) = \exp(-(t/\eta)^m)$ では $h(t) = (m/\eta)(t/\eta)^{m-1}$

$\eta^{(m-1)}$ で、 $m > 1$ なら増加(摩耗型)、 $m < 1$ なら減少(初期不良型)、 $m = 1$ なら一定(指数分布)です。

数値例です。故障までの時間が指数分布で $\lambda = 0.5$ (1時間あたり)のとき、3時間より長く動く確率は次の値です。

$$S(3) = e^{-(1.5)} = 0.2231$$

すでに3時間動いた装置がさらに3時間動く確率も、同じ 0.2231 です。これが指数分布の無記憶性で、一定のハザードと同じ内容です(経過時間は確率に影響しません)。

F と S の取り違えに注意してください。F は「t 以下」、S は「t より大きい」の確率で、足すと1です(例: $F(3) = 0.7769$)。

1.2-3 同時分布から周辺分布を出す(同時確率関数・同時確率密度関数)

2つの確率変数 X, Y の同時確率関数は $f(x, y) = P(X = x, Y = y)$ で、 $f \geq 0$ と $\sum \sum f = 1$ を満たします。同時確率密度関数は $f(x, y) \geq 0$ と $\iint f \, dx \, dy = 1$ を満たし、領域 D の確率は $\iint_D f \, dx \, dy$ です。周辺分布は、もう一方の変数を足し込むか、積分して消した分布です。

$$\text{離散: } f_x(x) = \sum_y f(x, y)$$

$$\text{連続: } f_x(x) = \int f(x, y) \, dy$$

独立の条件は、すべての (x, y) で $f(x, y) = f_x(x) f_y(y)$ となることです。離散の表なら、1マスずつ積を確認します。連続では、台が長方形でなければ独立になりません。

離散の例では、 $X \in \{0, 1, 2\}$ 、 $Y \in \{0, 1\}$ の同時確率が次の表で与えられるとします。

	Y=0	Y=1	周辺 f_x
X=0	0.1	0.2	0.3
X=1	0.2	0.1	0.3
X=2	0.1	0.3	0.4
周辺 f_y	0.4	0.6	1

$f(0,0) = 0.1$ に対して $f_x(0)f_y(0) = 0.3 \times 0.4 = 0.12$ で、一致しません。1マスでも不一致があれば、従属と言い切れます。

連続の例は、三角形の領域で $f(x, y) = 2$ ($0 < y < x < 1$) とした密度です。周辺密度は、積分範囲が x や y に依存します。

$$f_x(x) = \int_0^x 2 \, dy = 2x, \quad f_y(y) = \int_y^1 2 \, dx = 2(1 - y)$$

積 $4x(1 - y)$ は $f = 2$ に一致しないので、従属です。密度が定数でも、台が三角形なら積の形にはなりません。周辺分布から同時分布は復元できません。独立のときだけ、周辺の積が同時分布になります。

1.2-4 条件つき確率関数・条件つき確率密度関数

$X = x$ を固定したときの Y の分布が、条件つき分布です。

$$f(y|x) = f(x, y) / f_x(x) \quad (f_x(x) > 0)$$

y について足すか積分すると 1 になります。独立なら $f(y|x) = f_y(y)$ で、条件を付けても分布は変わりません。

前項の離散例で $X = 2$ のとき、次の値です。

$$f(Y=1|X=2) = 0.3 / 0.4 = 0.75$$

$X = 0$ なら $0.2/0.3 = 2/3$ 、 $X = 1$ なら $0.1/0.3 = 1/3$ となり、 X の値で条件つき分布が動きます。ここからも、従属と判断できます(独立なら全行で同じ分布になります)。

三角形の例では $f(y|x) = 2/(2x) = 1/x$ ($0 < y < x$) です。 $X = x$ を知ると、 Y は区間 $(0, x)$ 上の一様分布に従います。したがって、条件つき期待値と条件つき分散は次の値になります。

$$E[Y|X = x] = x/2, \quad V[Y|X = x] = x^2/12$$

割る相手は、条件にした変数の周辺分布です。同時分布どうしの比ではありません。二変量正規分布の条件つき分布が正規分布になる事実は、第2章で使います。

正方形 $(0, 1)^2$ 上の $f(x, y) = x + y$ では、 $f_x(x) = x + 1/2$ なので、条件つき密度は次の形です。

$$f(y|x) = (x + y)/(x + 1/2), \quad X = 1/2 \text{ のとき } f(y|x) = 1/2 + y$$

このとき $P(Y > 1/2 | X = 1/2) = \int (1/2 \text{ から } 1)(1/2 + y)dy = 0.25 + 0.375 = 0.625$ です。 $E[Y|X = 1/2]$ は $\int_0^1 y(1/2 + y)dy = 1/4 + 1/3 = 7/12$ になります。

1.2-5 モーメント母関数(定義・モーメントの取り出し・分布の特定)

モーメント母関数(積率母関数)は、0 の近傍で期待値が存在するときの $M(t) = E[e^{tX}]$ です。出題の中心は次の3つの性質です。

$$M^{(k)}(0) = E[X^k] \quad (k \text{ 階微分の } t = 0 \text{ での値})$$

母関数は分布を一意に決める

独立な X, Y について $M_{\{X+Y\}}(t) = M_X(t) M_Y(t)$

主な分布の母関数は次の表のとおりです。

分布	M(t)
二項分布 Bin(n, p)	$(1 - p + p e^t)^n$
ポアソン分布 Po(λ)	$\exp(\lambda(e^t - 1))$
指数分布 Exp(λ)	$\lambda / (\lambda - t) \ (t < \lambda)$
正規分布 N(μ, σ²)	$\exp(\mu t + \sigma^2 t^2 / 2)$

ポアソン分布でモーメントを取り出します。微分すると次の式が得られます。

$$M'(t) = \lambda e^t M(t), \quad M''(t) = (\lambda e^t + \lambda^2 e^{2t}) M(t)$$

$t = 0$ を代入して、 $E[X] = \lambda, E[X^2] = \lambda + \lambda^2$ です。分散は $E[X^2] - E[X]^2 = \lambda$ です (ポアソン分布は平均と分散が等しい)。指数分布 $M(t) = \lambda / (\lambda - t)$ では、 $M'(t) = \lambda / (\lambda - t)^2$ から $E[X] = 1/\lambda, M''(0) = 2/\lambda^2$ から $V[X] = 1/\lambda^2$ です。 $M(t) = (0.3 + 0.7e^t)^5$ は Bin(5, 0.7) と読めるので、 $E[X] = 3.5, V[X] = 1.05$ になります。

分布の特定は、母関数の形を表と見比べて行います。 $M(t) = 0.2 + 0.3e^t + 0.5e^{(2t)}$ なら、0, 1, 2 を確率 0.2, 0.3, 0.5 でとる離散分布です($e^{(kt)}$ の係

数が $P(X = k)$ 。 $M(t) = \exp(t + 2t^2)$ は、 $\mu t + \sigma^2 t^2 / 2$ との比較で $\mu = 1, \sigma^2 = 4$ の正規分布と決まります。独立な $Po(2)$ と $Po(3)$ の和は、母関数の積が $\exp(5(e^t - 1))$ になるので $Po(5)$ です。

1次変換では $M_{\{aX+b\}}(t) = e^{bt} M_X(at)$ を使います。母関数が存在しない分布もあります(コーシー分布)。

1.2-6 確率母関数(離散分布の母関数・独立な和の分布)

非負整数値の X には、確率母関数 $G(s) = E[s^X] = \sum P(X = k) s^k$ を定義できます。

$$G(1) = 1, \quad G'(1) = E[X], \quad V[X] = G''(1) + G'(1) - G'(1)^2$$

$$P(X = k) = G^{(k)}(0) / k! \quad (s^k \text{ の係数})$$

$$\text{独立な和: } G_{\{X+Y\}}(s) = G_X(s) G_Y(s)$$

$$\text{モーメント母関数との関係: } M(t) = G(e^t)$$

$G''(1)$ は $E[X(X - 1)]$ に当たり、 $E[X^2] = G''(1) + G'(1)$ から上の分散の式が出ます。

$G(s) = (0.2 + 0.8s)^3$ は、成功確率 0.8 の試行を3回行う $\text{Bin}(3, 0.8)$ です。 $E[X] = G'(1) = 3 \times 0.8 = 2.4$ 、 $V[X] = 3 \times 0.8 \times 0.2 = 0.48$ です。展開しても $P(X = 0) = 0.008$ のように確率を直接読めます (0.2^3)。ポアソン分布は $G(s) = \exp(\lambda(s - 1))$ で、 $G'(1) = \lambda$ 、 $G''(1) = \lambda^2$ なので、 $V[X] = \lambda^2 + \lambda - \lambda^2 = \lambda$ です。

サイコロ2個の目の和は、 $G(s) = ((s + s^2 + \dots + s^6)/6)^2$ の係数から読めます。 s^7 の係数は $6/36$ なので、 $P(\text{和} = 7) = 1/6$ です (36通り中6通り)。 $\text{Bin}(3, 0.8)$ と $\text{Bin}(2, 0.8)$ が独立なら、 G の積 $(0.2 + 0.8s)^5$ から、和は $\text{Bin}(5, 0.8)$ と分かります。

独立な和の分布は、母関数の積を作って係数を読めば、畳み込みの和を計算せずに出ます。使い分けは、整数値の離散分布で確率そのものを知りたいなら確率母関数、連続分布を含めて積率を知りたいならモーメント母関数です。

分布の特性値 [1.3]

この節の語は、次の9つです。

- モーメント、歪度、尖度、変動係数
- 相関係数、偏相関係数
- 分位点関数
- 条件つき期待値、条件つき分散

前半は1つの分布の形を数値で表す道具、後半は2変数以上の関係と、条件を付けた平均・散らばりを扱います。

1.3-1 期待値・分散・モーメント(原点・中心)の計算と線形性

原点まわりの k 次モーメントは $E[X^k]$ 、平均 μ まわり(中心)の k 次モーメントは $E[(X - \mu)^k]$ です。2次の中心モーメントが分散に当たります。標本版は、原点まわりなら $\Sigma x_i^k/n$ です。

$$E[aX + b] = aE[X] + b$$

$$V[X] = E[X^2] - E[X]^2$$

$$V[aX + b] = a^2V[X]$$

期待値の線形性は、独立かどうかに関係なく常に成り立ちます。分散の式で定数 b は消え、 a は2乗で効きます。一般に、 $aX + b$ の k 次の中心モーメントは X の k 次の中心モーメントの a^k 倍になります (b は消えます)。

中心モーメントは、原点まわりのモーメントから組み立てます。3次と4次の式は次のとおりです。

$$\mu_3 = E[X^3] - 3\mu E[X^2] + 2\mu^3$$

$$\mu_4 = E[X^4] - 4\mu E[X^3] + 6\mu^2 E[X^2] - 3\mu^4$$

例として、平均1の指数分布($\lambda = 1$)では $E[X^k] = k!$ なので、 $E[X] = 1$ 、 $E[X^2] = 2$ 、 $E[X^3] = 6$ 、 $E[X^4] = 24$ です。分散は $2 - 1 = 1$ 、 $\mu_3 = 6 - 6 + 2 = 2$ 、 $\mu_4 = 24 - 24 + 12 - 3 = 9$ になります。この μ_3 と μ_4 が、次項の歪度と尖度の材料になります ($\sigma = 1$ なので割り算は不要)。標本モーメント $\Sigma x_i^k / n$ を母モーメントに等しいと置く方法が、モーメント法です(第3章)。

間違えやすいのは $E[X^2]$ と $E[X]^2$ の混同です。 $E[X^2]$ は $E[X]^2$ 以上で、差が分散です。

1.3-2 歪度と尖度(定義・分布の形との対応)

標準偏差を σ として、2つの指標は次のように定義します。

$$\text{歪度} = E[(X - \mu)^3] / \sigma^3$$

$$\text{尖度} = E[(X - \mu)^4] / \sigma^4$$

歪度は左右の非対称性を表します。0 なら対称に近く、正なら右に裾が長く、負なら左に裾が長い分布です。左右対称な分布の歪度は 0 ですが、逆は成り立ちません。歪度 0 でも非対称な分布は作れます。

尖度は裾の重さに関係が深い指標です。正規分布の尖度は 3、一様分布は 1.8 です。尖度から 3 を引いた値を超過尖度(正規分布で 0)と呼び、この値を

尖度と呼ぶ流儀もあります。問題文の定義を最初に確かめなければ、答えが3ずれても気づけません。

分布	歪度	尖度	超過尖度
正規分布	0	3	0
一様分布	0	1.8	-1.2
指数分布	2	9	6
ポアソン分布 $Po(\lambda)$	$1/\sqrt{\lambda}$	$3 + 1/\lambda$	$1/\lambda$

指数分布の2と9は、前項の $\mu_3 = 2$ 、 $\mu_4 = 9$ を $\sigma = 1$ で割った値です。標本から計算する歪度・尖度には、不偏補正を入れる定義と入れない定義があり、ソフトウェアで値が違ふことがあります。

1次変換 $aX + b$ では、 $a > 0$ なら歪度も尖度も変わりません。 $a < 0$ だと歪度の符号だけが反転し、尖度は変わりません。尖度が大きいことは、山が鋭いというより裾が重い、という読み方が中心です(山の鋭さは副次的)。

1.3-3 変動係数(単位の違う量の散らばりの比較)

変動係数は、標準偏差を平均で割った $CV = \sigma/\mu$ です(標本では $s/\bar{x}$)。単位の無い相対的な散らばりを表し、平均が正の比例尺度のデータに使います。

A店の客単価が平均50、標準偏差10、B店が平均200、標準偏差30だとします。標準偏差はB店の方が大きい値です。変動係数は次のとおりです。

A: $10/50 = 0.20,$ B: $30/200 = 0.15$

相対的な散らばりはA店の方が大きいと分かります。単位が違う量(身長と体重など)の散らばりも、同じ物差しで比べられます。

単位の換算(cm から m など、定数倍)では変動係数は変わりません。定数を足す変換(摂氏からケルビンなど)では平均だけが動くので、変動係数は変わります。このため、間隔尺度のデータには向きません。平均が 0 に近いと値が不安定になるため、平均が 0 付近の量にも使えません。参考として、指数分布の CV は 1、ポアソン分布は $1/\sqrt{\lambda}$ です。

1.3-4 共分散と相関係数(性質・独立と無相関の違い)

共分散と相関係数は次の式で定義します。

$$\text{Cov}(X, Y) = E[(X - \mu_x)(Y - \mu_y)] = E[XY] - E[X]E[Y]$$

$$\rho = \text{Cov}(X, Y) / \sqrt{V[X]V[Y]}, \quad -1 \leq \rho \leq 1$$

$\rho = \pm 1$ になるのは、 $Y = aX + b$ ($a \neq 0$) と完全な直線関係にあるときだけです。相関係数が測るのは線形関係の強さで、曲線的な関係は測れません。

変換に対する性質は次のとおりです。

$$\text{Cov}(aX + b, cY + d) = ac \text{Cov}(X, Y)$$

$$V[aX + bY] = a^2V[X] + b^2V[Y] + 2ab \text{Cov}(X, Y)$$

相関係数は、位置と尺度の変更($ac > 0$)で変わりません。 $ac < 0$ なら符号だけが変わります。

数値例では、 $V[X] = 4$ 、 $V[Y] = 9$ 、 $\rho = 0.5$ のとき $\text{Cov} = 0.5 \times 2 \times 3 = 3$ です。 $V[X - 2Y]$ は次の値になります。

$$4 + 4 \times 9 + 2 \times 1 \times (-2) \times 3 = 4 + 36 - 12 = 28$$

独立なら共分散の項が消えて 40 です。

独立なら共分散は 0 で、相関係数も 0 です。逆は成り立ちません。 X が区間 $(-1, 1)$ の一様分布で $Y = X^2$ のとき、 $\text{Cov}(X, Y) = E[X^3] - E[X]E[X^2] = 0$ ですが、 Y は X の関数で独立ではありません。二変量正規分布に限っては、相関係数 0 と独立が同値です。

前節の三角形の密度($f = 2, 0 < y < x < 1$)では、 $E[XY] = 1/4$ 、 $E[X] = 2/3$ 、 $E[Y] = 1/3$ 、 $V[X] = V[Y] = 1/18$ から、次の値が出ます。

$$\text{Cov} = 1/4 - 2/9 = 1/36, \quad \rho = (1/36)/(1/18) = 0.5$$

1.3-5 偏相関係数(第3の変数の影響を除く・相関係数からの計算)

変数 Z の影響を除いたときの X と Y の相関が、偏相関係数です。単相関の3つから計算できます。

$$\rho_{XY \cdot Z} = (\rho_{XY} - \rho_{XZ} \rho_{YZ}) / \sqrt{(1 - \rho_{XZ}^2)(1 - \rho_{YZ}^2)}$$

この値は、 X を Z で回帰した残差と、 Y を Z で回帰した残差の相関に等しくなります。 Z の線形な効果を取り除いたあとに残る関係を見る量です。

数値例です。 $\rho_{XY} = 0.6$ 、 $\rho_{XZ} = 0.5$ 、 $\rho_{YZ} = 0.4$ のとき、次の値です。

$$(0.6 - 0.5 \times 0.4) / \sqrt{(0.75 \times 0.84)} = 0.4 / 0.7937 = 0.504$$

単相関 0.6 より小さくなったのは、 Z を共通の原因として一部の相関を説明できたからです。

符号が反転する場合もあります。 $\rho_{XY} = 0.5$ 、 $\rho_{XZ} = \rho_{YZ} = 0.8$ なら、 $(0.5 - 0.64)/(1 - 0.64) = -0.389$ です。単相関が正でも、 Z を固定すると負になります。 $\rho_{XY} = \rho_{XZ} = \rho_{YZ}$ なら偏相関係数は 0 で、 X と Y の相関は Z だけで説明できます(見かけの相関)。

間違えやすいのは、単相関と偏相関を取り違えることです。3つの単相関がそろわないと計算できません。重回帰分析の偏回帰係数とは別の量なので、第6章で区別します。

1.3-6 分位点関数(累積分布関数の逆・分位点の計算)

分位点関数は、累積分布関数 F の逆の働きをする関数です。

$$Q(p) = F^{-1}(p) = \inf\{x : F(x) \geq p\}, \quad 0 < p < 1$$

$Q(p)$ を p 分位点といい、中央値は $Q(0.5)$ 、四分位点は $Q(0.25)$ 、 $Q(0.75)$ です。 F が連続で狭義単調増加なら、 $F(x) = p$ を解いた x が $Q(p)$ です。標準正規分布の上側 5% 点 1.645 は、 $Q(0.95)$ に当たります(上側確率 0.05 の点)。

$f(x) = 2x$ ($0 \leq x \leq 1$) なら $F(x) = x^2$ なので、 $Q(p) = \sqrt{p}$ です。

$$Q(0.25) = 0.5, \quad Q(0.5) = 0.7071, \quad Q(0.75) = 0.8660$$

指数分布 $\text{Exp}(\lambda)$ では $Q(p) = -\ln(1 - p)/\lambda$ で、中央値は $\ln 2/\lambda$ です。 $\lambda = 2$ の中央値は $0.6931/2 = 0.3466$ になります($\ln 2 = 0.6931$)。

離散分布では F が階段になるので、定義の $\inf$ の意味が効きます。 $X = 0, 1, 2$ が確率 0.3, 0.3, 0.4 のとき、 F は 0.3, 0.6, 1 です。次の値になります。

$$Q(0.3) = 0, \quad Q(0.5) = 1, \quad Q(0.6) = 1, \quad Q(0.61) = 2$$

p がちょうど F の値に一致する点 ($Q(0.3) = 0$ 、 $Q(0.6) = 1$) では、 $F(x) \geq p$ を初めて満たす x 、つまり F が p を超える点より 1 つ前の値になります。ここが間違えやすい所です。

分位点関数は、乱数生成の逆関数法の土台です。 U が区間 $(0, 1)$ の一様分布に従うとき、 $Q(U)$ の分布は F になります(逆関数法)。 $\lambda = 2$ の指数分布で $U = 0.75$ なら、次の乱数が得られます。

$$Q(0.75) = \ln 4/2 = 1.3863/2 = 0.6931$$

1.3-7 条件つき期待値と条件つき分散(繰り返し期待値の法則・全分散の公式)

$X = x$ を与えたときの Y の期待値 $E[Y|X = x]$ は x の関数です。 X を確率変数のままにした $E[Y|X]$ は、それ自体が確率変数になります。条件つき分散も同様です(次の式)。

繰り返し期待値の法則: $E[E[Y|X]] = E[Y]$

$$V[Y|X] = E[Y^2|X] - (E[Y|X])^2$$

全分散の公式: $V[Y] = E[V[Y|X]] + V[E[Y|X]]$

$E[Y|X = x]$ は回帰関数に当たり、 x で Y を予測するときの平均的な値を与えます。全分散の公式は、散らばりを「条件の中の散らばりの平均」と「条件ごとの平均の散らばり」に分けます。群内分散と群間分散に分ける発想と同じ構造で、階層構造(学校の中の生徒など)の分散分解に使います(第7章の分散分析)。

三角形の密度では、条件つき分布が $(0, x)$ 上の一様分布なので、 $E[Y|X] = X/2$ 、 $V[Y|X] = X^2/12$ です。 $E[X] = 2/3$ 、 $E[X^2] = 1/2$ 、 $V[X] = 1/18$ を使うと、次の

ように計算できます。

$$E[Y] = E[X]/2 = 1/3$$

$$V[Y] = E[X^2]/12 + V[X]/4 = 1/24 + 1/72 = 1/18$$

$E[Y]$ と $V[Y]$ は、周辺密度 $2(1-y)$ から直接求めた値に一致します。

もう1つは、確率的な件数の合計です。1日の請求件数 N がポアソン分布 $Po(3)$ に従い、1件の金額が平均 10、分散 4 で独立だとします。合計 S について、 $E[S|N] = 10N$ 、 $V[S|N] = 4N$ です。

$$E[S] = 10 \times 3 = 30$$

$$V[S] = E[4N] + V[10N] = 12 + 100 \times 3 = 312$$

第1項の 12 だけで終わると、件数のばらつき(300)が抜けます。全分散の第2項を忘れる誤りは頻出です。離散の表の例(1.2-3)でも、 $V[Y] = 0.24$ は $E[V[Y|X]] = 0.2083$ と $V[E[Y|X]] = 0.0317$ の和に一致します($E[Y] = 0.6$)。

変数変換 [1.4]

この節の語は、変数変換、確率変数の線形結合の分布、Box-Cox変換の3つです。確率変数に関数を施したときの分布(密度)の求め方と、線形結合の平均・分散・分布、データを正規分布に近づける冪変換を扱います。

1.4-1 1変数の変数変換(ヤコビアン・単調変換の密度)

連続確率変数 X の密度を f_X とし、 $Y = g(X)$ とします。 g が単調で逆関数 $x = g^{-1}(y)$ をもつなら、 Y の密度は次の式です。

$$f_Y(y) = f_X(g^{-1}(y)) |dx/dy|$$

絶対値を付けるのは、 g が減少でも密度が負にならないようにするためです。 Y のとりうる範囲(台)も、 X の台を g で写した範囲へ変わります(出発点は X の台)。

X が区間 $(0, 1)$ の一様分布($f_X = 1$)のとき、 $Y = -\ln X$ では $x = e^{-y}$ で $|dx/dy| = e^{-y}$ なので、 $f_Y(y) = e^{-y}$ ($y > 0$) です。平均1の指数分布が得られます(逆関数法の例)。 $Y = X^2$ では $x = \sqrt{y}$ で $|dx/dy| = 1/(2\sqrt{y})$ なので、 $f_Y(y) = 1/(2\sqrt{y})$ ($0 < y < 1$) です。

g が単調でないときは、区間ごとに分けて足します。 X が標準正規分布のとき、 $Y = X^2$ は $y = x^2$ の根 $x = \pm\sqrt{y}$ を2本もつので、次の密度になります。

$$f_Y(y) = 2 \times \phi(\sqrt{y}) \times 1/(2\sqrt{y}) = e^{-y/2} / \sqrt{2\pi y} \quad (y > 0)$$

これは自由度1のカイ二乗分布の密度です。もう1つの方法は累積分布関数法で、 $P(Y \leq y) = P(-\sqrt{y} \leq X \leq \sqrt{y})$ と書き、 y で微分して密度を得ます。単調でない変換では、こちらの方が間違えにくい手順です。

1次変換と対数正規分布の密度も、同じ式から出ます。

$$Y = aX + b \ (a \neq 0): \ f_Y(y) = f_X((y - b)/a) / |a|$$

$$X \sim N(\mu, \sigma^2), \ Y = e^X: \ f_Y(y) = \phi((\ln y - \mu)/\sigma) / (\sigma y) \quad (y > 0)$$

後者が対数正規分布の密度で、 $\ln Y$ が正規分布に従う分布です。

1.4-2 2変数の変数変換(和・比・積の分布)

(X, Y) から $(U, V) = (g_1(X, Y), g_2(X, Y))$ への変換では、逆変換 $x = x(u, v)$ 、 $y = y(u, v)$ を求め、ヤコビアン¹の絶対値を掛けます。

$$f_{\{U,V\}}(u, v) = f_{\{X,Y\}}(x(u, v), y(u, v)) |J|, \quad J = \partial(x, y) / \partial(u, v)$$

X, Y が独立な標準正規分布に従うとき、 $U = X + Y, V = X - Y$ を考えます。逆変換は $x = (u + v)/2, y = (u - v)/2$ で、 $J = (1/2)(-1/2) - (1/2)(1/2) = -1/2$ より $|J| = 1/2$ です。 $x^2 + y^2 = (u^2 + v^2)/2$ を使って整理すると、次の式になります。

$$f_{\{U,V\}}(u, v) = (1/4\pi) \exp(-(u^2 + v^2)/4)$$

これは $N(0, 2)$ の密度2つの積なので、 U と V は独立で、どちらも $N(0, 2)$ に従います。

独立な $\text{Exp}(1)$ の X, Y から、 $U = X + Y, V = X/(X + Y)$ を作る例も定番です。逆変換は $x = uv, y = u(1 - v)$ で、ヤコビアン¹の絶対値は $|J| = u$ です。

$$f_{\{U,V\}}(u, v) = e^{-x} e^{-y} \times u = u e^{-u} \quad (u > 0, 0 < v < 1)$$

密度が u の関数と v の関数の積に分かれるので、 U と V は独立で、 U はガンマ分布(形状 2)、 V は区間 $(0, 1)$ の一様分布です。

和の分布は、畳み込み $f_Z(z) = \int f_X(x) f_Y(z - x) dx$ で求められます。独立な $\text{Exp}(1)$ を2つ足すと、 $f_Z(z) = z e^{-z} (z > 0)$ になります(形状 2 のガンマ分布)。形状 2 のガンマ分布 $\text{Ga}(2, \lambda)$ の上側確率は $P(T > t) = e^{-\lambda t}(1 + \lambda t)$ で、率 λ のポアソン過程で時刻 t までの件数が 1 件以下である確率に当たります。比 X/Y (独立な標準正規分布どうし)はコーシー分布に従います。積は、

対数をとれば和の問題に直せます($\ln(XY) = \ln X + \ln Y$)。最大値や最小値は、変換ではなく累積分布関数から $P(\max \leq x) = F(x)^n$ のように求めます。

1.4-3 確率変数の線形結合の期待値・分散・分布(独立な正規の和)

線形結合 $\sum a_i X_i$ の期待値と分散は、次のとおりです。

$$E[\sum a_i X_i] = \sum a_i E[X_i] \quad (\text{常に成立})$$

$$V[\sum a_i X_i] = \sum a_i^2 V[X_i] + 2 \sum_{\{i < j\}} a_i a_j \text{Cov}(X_i, X_j)$$

独立なら共分散の項が消えます。分布についての性質は、次の4つの再生性です。

- 独立な正規分布の線形結合 → 正規分布
- 独立なポアソン分布の和 → ポアソン分布(母数は和)
- 尺度が同じ独立なガンマ分布の和 → ガンマ分布(形状は和)
- 独立なカイ二乗分布の和 → カイ二乗分布(自由度は和)

特に、 $X_1, \dots, X_n$ が独立で $N(\mu, \sigma^2)$ に従うとき、標本平均は $N(\mu, \sigma^2/n)$ に従います。

数値例です。 X が $N(50, 10^2)$ 、 Y が $N(60, 8^2)$ で独立のとき、 X が Y を上回る確率を求めます。 $D = X - Y$ は平均 -10 、分散 $100 + 64 = 164$ の正規分布で、標準偏差は 12.81 です。

$$P(X > Y) = P(D > 0) = P(Z > 10/12.81) = P(Z > 0.78) = 0.218$$

平均が低い X が上回る確率は、2割ほどです。引き算でも、分散は引かずに足さなければなりません。

間違えやすいのは $2X$ と $X_1 + X_2$ の区別です。同じ変数を2倍した $2X$ の分散は $4\sigma^2$ ですが、独立な2つのコピーの和 $X_1 + X_2$ の分散は $2\sigma^2$ にとどまります。

1.4-4 Box-Cox 変換(目的・ λ と対数変換)

$x > 0$ のデータに対する Box-Cox 変換は、次のとおりです。

$$y(\lambda) = (x^\lambda - 1)/\lambda \quad (\lambda \neq 0), \quad y(0) = \ln x$$

$\lambda \rightarrow 0$ の極限が対数なので、 $\lambda = 0$ で対数変換になります。代表的な値は次の表のとおりです($\lambda = 1$ は実質的な変換なし)。

λ	変換	意味
1	$x - 1$	平行移動だけ(実質的な変換なし)
0.5	$2(\sqrt{x} - 1)$	平方根変換
0	$\ln x$	対数変換
-1	$1 - 1/x$	逆数変換

目的は、分布を正規分布に近づけ、分散を一定に近づけ、関係を線形に近づけることです。右に裾が長いデータは、 $\lambda < 1$ の方向で対称に近づきます(平方根や対数)。 λ は最尤法で推定します(正規分布を仮定した尤度が最大になる λ を探す)。

$x > 0$ が必要なので、0 や負の値を含むデータには、定数を足してから使います。変換後の尺度で推測し、結果は逆変換で戻します。単調変換なので中央値は逆変換で保たれますが、平均は保たれません。

極限定理, 漸近理論 [1.5]

この節の語は、次の13個です。

- 概収束、確率収束、分布収束
- 大数の弱法則、少数法則、中心極限定理
- 二項分布の正規近似、ポアソン分布の正規近似、連続修正
- 極値分布
- 連続写像定理、スルツキーの補題、デルタ法

推測統計の理屈は、標本サイズ n が大きいときの振る舞いに支えられています。この節は、その振る舞いを3種類の収束で区別し、近似の使い分けと、推定量の分布の求め方までを扱います。

1.5-1 収束の種類(概収束・確率収束・分布収束)と強弱の関係

確率変数の列 X_n が X に近づくことを、3通りに定義します。

収束	定義
概収束	$P(\lim X_n = X) = 1$ (ほとんど確実に収束)
確率収束	任意の $\epsilon > 0$ で $P(X_n - X > \epsilon) \rightarrow 0$
分布収束	$F_n(x) \rightarrow F(x)$ (F の連続点で)

強弱の関係は、次の一方向です。

概収束 ⇒ 確率収束 ⇒ 分布収束

逆は一般に成り立ちません。ただし、定数への分布収束は確率収束と同値です。大数の強法則は概収束、弱法則は確率収束、中心極限定理は分布収束の

主張で、一致性(推定量が母数に確率収束すること)は確率収束の言葉で書かれます。

確率収束しても期待値が収束しない例があります。 X_n が確率 $1/n$ で n 、確率 $1 - 1/n$ で 0 をとる列では、 $P(X_n > \varepsilon) = 1/n \rightarrow 0$ なので 0 に確率収束しますが、 $E[X_n] = 1$ は 0 に近づきません。確率収束は、期待値の収束を保証しません。ただし、 $|X_n| \leq M$ となる定数 M があれば(有界)、確率収束から $E[X_n] \rightarrow E[X]$ が言えます(有界収束定理)。上の反例の列は n とともに値が大きくなるので、有界ではありません。

場面での判断は、次の対応で行います。 X_n が $\text{Bin}(n, \lambda/n)$ のとき、 $\text{Po}(\lambda)$ への収束は分布収束で、確率変数どうしの近さ(確率収束)は主張していません。標本平均 $\bar{x}_n$ が μ に近づくのは確率収束(弱法則)または概収束(強法則)で、 $\sqrt{n}(\bar{x}_n - \mu)/\sigma$ が $N(0, 1)$ に近づくのは分布収束です。

分布収束しても確率収束しない例もあります。 X が $N(0, 1)$ に従い、 X_n を n が偶数なら X 、奇数なら $-X$ とします。どの X_n も $N(0, 1)$ なので $N(0, 1)$ に分布収束しますが、奇数の n では $|X_n - X| = 2|X|$ が小さくならず、 X に確率収束しません。

1.5-2 大数の弱法則(チェビシェフの不等式による評価)

$X_1, \dots, X_n$ が独立で同じ分布に従い、平均 μ をもつなら、標本平均は μ に確率収束します。これが大数の弱法則で、概収束を主張するのが強法則です。有限の分散があれば、チェビシェフの不等式で簡単に示せます。

マルコフの不等式: $P(Y \geq a) \leq E[Y]/a \quad (Y \geq 0, a > 0)$

チェビシェフの不等式: $P(|X - \mu| \geq \varepsilon) \leq \sigma^2 / \varepsilon^2$

標本平均に当てはめると、 $V[\bar{x}] = \sigma^2/n$ なので、次の評価になります。

$$P(|\bar{X} - \mu| \geq \varepsilon) \leq \sigma^2 / (n \varepsilon^2)$$

右辺は $n \rightarrow \infty$ で 0 に向かいます(弱法則の証明)。k 倍の標準偏差で書けば $P(|\bar{X} - \mu| \geq k\sigma) \leq 1/k^2$ で、k=2 なら 0.25、k=3 なら 0.111 が上限です。

標本サイズの見積もりに使います。 $\sigma = 2$ の母集団で、 $P(|\bar{x} - \mu| < 0.5) \geq 0.95$ を保証したいとします。余事象の確率を 0.05 以下にすればよいので、次の不等式を解きます。

$$\sigma^2 / (n \varepsilon^2) = 4 / (n \times 0.25) \leq 0.05 \quad \rightarrow \quad n \geq 320$$

中心極限定理で正規近似すると、 $1.96 \times 2 / 0.5 = 7.84$ の 2 乗から $n \geq 61.5 (7.84^2 = 61.47)$ 、つまり $n = 62$ で足ります。チェビシェフの評価は、分布の形を仮定しない代わりに、大きすぎる n を要求します。

不等式は上からの評価であり、等号の保証ではありません。弱法則は平均の存在が前提で、コーシー分布のように平均がない分布では成り立ちません(標本平均が再びコーシー分布に従います)。

1.5-3 中心極限定理(標本平均の近似分布)

$X_1, \dots, X_n$ が独立で同じ分布に従い、平均 μ 、有限の分散 σ^2 をもつとき、標準化した標本平均は標準正規分布に分布収束します。

$$\sqrt{n} (\bar{X} - \mu) / \sigma \rightarrow N(0, 1) \quad (n \rightarrow \infty)$$

母集団の分布の形は問いません。実用では、 $\bar{x}$ が近似的に $N(\mu, \sigma^2/n)$ 、合計 S_n が近似的に $N(n\mu, n\sigma^2)$ に従うと読みます。 σ が未知なら、標本の標準偏差 s で置き換えられます(1.5-7)。

サイコロを100回振った目の合計 S の例です。1回の平均は 3.5、分散は $35/12 = 2.9167$ なので、 S の平均は 350、分散は 291.67、標準偏差は 17.08 です。 $P(S > 370)$ を正規近似すると、次のとおりです。

$$z = (370 - 350)/17.08 = 1.17, \quad P(Z > 1.17) = 0.121$$

標本平均の確率にも使えます。母集団の平均 50、標準偏差 12 から $n = 64$ の標本をとると、 $\bar{x}$ の標準偏差は $12/\sqrt{64} = 1.5$ です。

$$P(\bar{x} > 52) = P(Z > 2/1.5) = P(Z > 1.33) = 0.092$$

連続修正を入れる精密な形は、次項で扱います。

収束の速さは、有限の n では保証されません。母集団が強く歪んでいると、 n が小さいうちは近似が悪くなります。分散が無限の分布(コーシー分布など)では定理が成り立ちません。「 n が大きければ標本のヒストグラムが正規分布に近づく」という読み違いにも注意してください。近づくのは標本平均の分布であり、標本そのものの分布ではありません。

1.5-4 二項分布・ポアソン分布の正規近似と連続修正

二項分布 $\text{Bin}(n, p)$ は、 n が大きいとき近似的に $N(np, np(1-p))$ に従います。目安は $np \geq 5$ かつ $n(1-p) \geq 5$ です(10 を基準にする流儀もあります)。ポアソン分布 $\text{Po}(\lambda)$ は、 λ が大きい(目安 $\lambda \geq 10$)とき近似的に $N(\lambda, \lambda)$ に従います。 $\text{Po}(\lambda)$ は $\text{Po}(\lambda/n)$ の n 個の和なので、中心極限定理から導けます。

離散分布を連続分布で近似するときは、端点に ± 0.5 の補正をします(連続修正)。

$$P(X \leq k) \approx \Phi((k + 0.5 - \mu)/\sigma)$$

$$P(X \geq k) \approx 1 - \Phi((k - 0.5 - \mu)/\sigma)$$

$$P(X = k) \approx \Phi((k + 0.5 - \mu)/\sigma) - \Phi((k - 0.5 - \mu)/\sigma)$$

不等号の向きに迷ったら、整数点での「以下」「以上」の形に直してから補正します。

$$P(X < k) = P(X \leq k - 1) \rightarrow \text{境界は } k - 0.5$$

$$P(X > k) = P(X \geq k + 1) \rightarrow \text{境界は } k + 0.5$$

例 1: X が $\text{Bin}(100, 0.4)$ に従うとき、 $P(X \leq 45)$ を求めます。 $\mu = 40$ 、 $\sigma^2 = 24$ 、 $\sigma = 4.899$ です。連続修正ありでは、次のとおりです。

$$z = (45.5 - 40)/4.899 = 1.12, \quad P \approx 1 - 0.1314 = 0.869$$

厳密な値は 0.8689 です。修正なしだと $z = 1.02$ 、上側確率 0.1539 から 0.846 となり、精度が落ちます。

例 2: X が $\text{Po}(16)$ に従うとき、 $P(X \geq 20)$ を求めます。 $\mu = 16$ 、 $\sigma = 4$ で、次の値になります。

$$z = (19.5 - 16)/4 = 0.875 \approx 0.88, \quad P \approx 0.189$$

厳密な値は 0.188 です。修正なしなら $z = 1$ で 0.159 となります。

標本比率 X/n も同じ形で、近似的に $N(p, p(1-p)/n)$ に従い、連続修正は $0.5/n$ の補正です(n で割った形)。

1.5-5 少数法則(二項分布のポアソン近似)

$\text{Bin}(n, p)$ で $np = \lambda$ を一定に保ったまま $n \rightarrow \infty$ とすると、確率関数はポアソン分布 $\text{Po}(\lambda)$ の確率関数に収束します。これが少数の法則(ポアソンの少数法則)で、 n が大きく p が小さいときに二項分布をポアソン分布で近似する根拠です。「大数の法則」とは名前が似ていても、同じ定理ではありません。

数値例では、 $\text{Bin}(200, 0.01)$ は $\lambda = np = 2$ で、 $\text{Po}(2)$ に近づきます。

$P(X = 0):$

厳密 $0.99^{200} = 0.1340,$

近似 $e^{-2} = 0.1353$

$P(X = 2):$

厳密 $0.2720,$

近似 $2e^{-2} = 0.2707$

近似の使い分けは、 $np(1 - p)$ が大きければ正規近似、 p が小さく np が小さい〜中程度ならポアソン近似です。 $n = 1000, p = 0.002 (\lambda = 2)$ はポアソン近似、 $n = 100, p = 0.4$ は正規近似を選びます(分散 24)。

1.5-6 極値分布(最大値・最小値の極限分布)

独立で同じ分布 F に従う $X_1, \dots, X_n$ の最大値 M_n について、次が成り立ちます。

$$P(M_n \leq x) = F(x)^n$$

調整しないと、 $F(x) < 1$ の x では $F(x)^n \rightarrow 0$ となり、 M_n は分布の上端点に集まります(一点分布に退化)。適切に位置と尺度を調整した M_n の極限は、一般化極値分布の3つの型のどれかになります。

型	累積分布関数	元の分布の裾
Gumbel 型	$\exp(-e^{(-x)})$	指数的に減る裾(指数分布・正規分布)

型	累積分布関数	元の分布の裾
Fréchet 型	$\exp(-x^{-(\alpha)}) \ (x > 0)$	べき乗で減る重い裾(パレート分布)
Weibull 型	$\exp(-(-x)^{\alpha}) \ (x < 0)$	上限のある分布(一様分布)

3つは、一般化極値分布 $\exp(-(1 + \xi x)^{-1/\xi})$ に対応します ($\xi = 0$ が Gumbel、 $\xi > 0$ が Fréchet、 $\xi < 0$ が Weibull)。

指数分布 $\text{Exp}(1)$ の最大値は Gumbel 型です。

$$P(M_n - \ln n \leq x) = (1 - e^{-(x)/n})^n \rightarrow \exp(-e^{-(x)})$$

T 年に 1 度の値(再現レベル)は、年最大値の分布の累積確率 $1 - 1/T$ の点です(100 年に 1 度なら 99% 点)。

$n = 50$ 、 $x = 1$ のとき、厳密な値は 0.691、極限は 0.692 です。区間 $(0, 1)$ の一様分布では、 $n(1 - M_n)$ が $\text{Exp}(1)$ に分布収束し、上限のある Weibull 型になります。パレート分布 $P(X > x) = x^{-(\alpha)} \ (x \geq 1)$ では、 $M_n/n^{1/\alpha}$ が Fréchet 型に収束します。

最小値は、符号を反転した最大値として扱います($-X$ の最大値)。 $\text{Exp}(\lambda)$ の n 個の最小値は、ちょうど $\text{Exp}(n\lambda)$ に従います。年最大降水量のように、期間ごとの最大値(ブロック最大値)をあてはめるのが典型的な使い方です。

1.5-7 連続写像定理とスルツキーの補題

連続写像定理は、収束が連続関数を通して保たれる、という主張です。 X_n が X に分布収束し、 g が連続(または、 g の不連続点の集合に X が入る確率が 0) なら、 $g(X_n)$ は $g(X)$ に分布収束します。確率収束・概収束でも同じです(g が連続なら)。 $\bar{x}$ が μ に確率収束すれば、 $\bar{x}^2$ は μ^2 に、 $1/\bar{x}$ は $1/\mu \ (\mu \neq 0)$ に確率収束します。

束します。標準化した標本平均 Z_n が $N(0, 1)$ に分布収束すれば、 Z_n^2 は自由度1のカイ二乗分布に分布収束します。

スルツキーの補題は、分布収束する列と、定数に確率収束する列の組み合わせを扱います。 X_n が X に分布収束し、 Y_n が定数 c に確率収束するとき、次が成り立ちます。

$$X_n + Y_n \rightarrow X + c, \quad X_n Y_n \rightarrow cX, \quad X_n / Y_n \rightarrow X/c \quad (c \neq 0)$$

応用は、母標準偏差を標本標準偏差に置き換える場面です。 s が σ に確率収束するので、 σ の代わりに s を使っても極限は変わりません。

$$\sqrt{n} (\bar{X} - \mu) / s = [\sqrt{n} (\bar{X} - \mu) / \sigma] \times [\sigma / s] \rightarrow N(0, 1) \times 1$$

条件に注意してください。 Y_n は定数に収束しなければなりません。 Y_n が定数でない確率変数に収束するときは、結合分布の情報がないと和や積の極限が決まりません。 $X_n = Z, Y_n = -Z$ (Z は $N(0, 1)$) はどちらも $N(0, 1)$ に分布収束しますが、和は常に 0 です。独立な2つの $N(0, 1)$ の和は $N(0, 2)$ なので、同じ周辺分布でも結果が変わります。

1.5-8 デルタ法(関数の漸近分散)

推定量 T_n が $\sqrt{n} (T_n - \theta) \rightarrow N(0, \sigma^2)$ を満たし、 g が微分可能で $g'(\theta) \neq 0$ なら、 $g(T_n)$ の漸近分布は次のとおりです。

$$\sqrt{n} (g(T_n) - g(\theta)) \rightarrow N(0, g'(\theta)^2 \sigma^2)$$

$g(T_n)$ を θ のまわりで1次のテイラー展開すると得られる結果です。 $g'(\theta) = 0$ のときは、1次の項が消えるので、この式は使えません。

例 1(標本比率の対数オッズ): $p\hat{1}$ は $\sqrt{n}(\hat{p} - p) \rightarrow N(0, p(1 - p))$ を満たします。 $g(p) = \ln(p/(1 - p))$ の微分は $g'(p) = 1/(p(1 - p))$ なので、漸近分散は次のとおりです。

$$g'(p)^2 p(1 - p) = 1/(p(1 - p))$$

$p = 0.2, n = 400$ なら、 $g(p)$ の分散と標準偏差は次の値です。

$$1/(400 \times 0.16) = 0.015625, \quad \text{標準偏差 } 0.125$$

95% の区間幅は $\pm 1.96 \times 0.125 = \pm 0.245$ になります。

例 2(標本平均の2乗): $\mu = 3, \sigma = 2, n = 100$ で $g(x) = x^2$ とすると、 $g'(3) = 6$ です。 x^2 の分散は $6^2 \times 4/100 = 1.44$ です(標準偏差は 1.2)。

例 3(分散安定化): X がポアソン分布 $Po(\lambda)$ に従い、 λ が大きいとき、 $g(x) = \sqrt{x}$ の微分は $g'(\lambda) = 1/(2\sqrt{\lambda})$ です。分散は $\lambda \times 1/(4\lambda) = 1/4$ となり、 λ に依存しません。 $\sqrt{X}$ が近似的に $N(\sqrt{\lambda}, 1/4)$ に従うため、平均によって分散が変わるデータの分散を一定にする変換として使われます。

どの例も、手順は同じです。

- ①推定量の漸近分散を確認する。
- ② $g'(\theta)$ を求める。
- ③漸近分散に $g'(\theta)^2$ を掛ける。

平均の近似は $g(\theta)$ で、 θ が未知なら推定値で置き換えます(スルツキーの補題が根拠)。

サンプルはここまでです

続き(第2章～巻末)は、統計検定準1級 問題集のご購入で、頻出度別ノート・暗記用ノートと合わせて3点すべてダウンロードできます。

SAMPLE

統計検定準1級 学習用テキスト(無料サンプル)

版

2026年10月版(統計検定 公式「統計検定準1級 出題範囲表」(2027年1月1日改訂版・公式ページ表記 202107版) 準拠)

発行

IT資格道場(<https://www.shikaku-doujou.com>)

お問い合わせ

info@shikaku-doujou.com

「統計検定」は一般財団法人 統計質保証推進協会の登録商標です。記載されているその他の会社名・製品名・サービス名は、各社の商標または登録商標です。

本教材は公式シラバスの内容を基に作成したもので、一般財団法人 統計質保証推進協会および日本統計学会との提携・承認関係はありません。

本文・図表・設問はすべてオリジナルの著作物です。無断での転載・複製・再配布・二次販売はできません。お気づきの点は info@shikaku-doujou.com までご連絡ください。