

基礎統計（倉田博史） シケプリ（公式集）

理一 24 組（2010 年度入学）基礎統計シケ対

0 まえがき

0.1 はじめに

このシケプリは、基礎統計の講義、及び教科書に出てくる定義や定理をまとめた「公式集」です。

ここに載せた公式は、出来る限り理解・暗記したほうがいいと思われます。

あくまで「公式集」なので、ちゃんと勉強する時は教科書や他のクラスのシケプリを読んでください。（理二三 17 組のページに上がっているノートは非常によくまとまっていると思ひます）。

0.2 試験について

まず、講義プリントにある試験についての注意を載せておきます。

試験には電卓を持参のこと。それ以外の持ち込みは不可。

電卓は平方根が計算出来れば十分。安価なものでよい。

計算式や考え方が正しければ良く、計算精度を採点の対象とはしない。

関数電卓しか持っていない者はそれを使ってもよいが、関数計算機能など通常の電卓を超える機能を用いてはならない。

必ず出題するのは、

- 正規分布の確率計算
- 正規母集団、2 項母集団に関する推測問題

数表は配布する。e は 2.7 で計算する。試験時間は 90 分。

とのこと。ということで公式は暗記する必要があります。

問題傾向ですが、過去問を見る限り大半が計算問題です。たまに記述問題もあります（定義を述べよ、結果についてコメントせよ、など）。

分量はそこまで多くないと思われますが、計算練習はしておいた方がいいと思ひます。

1 データの整理と要約

1.1 1次元データ

1.1.1 平均 \bar{x}

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

1.1.2 平均は最小 2 乗値

$$\sum_{i=1}^n (c - x_i)^2 \text{ は } c = \bar{x} \text{ のとき最小.}$$

$$\sum_{i=1}^n (c - x_i)^2 \text{ は } c \text{ とデータ } x_1, x_2, \dots, x_n \text{ との「近さ」を表す.}$$

平均はデータに最も近い値である, ということ.

1.1.3 メディアン (中央値) Md

データ x_1, x_2, \dots, x_n を大きさの順に並べたものを $x_{(1)}, x_{(2)}, \dots, x_{(n)}$ とすると,

$$Md = \begin{cases} x_{(\frac{n+1}{2})} & (n: \text{奇数}) \\ \frac{x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}}{2} & (n: \text{偶数}) \end{cases}$$

データを大きさの順に並べたとき, その中央に来るものを「メディアン」という.

「中央値」ともいうので, “「メディアン」=「中央値」と覚えておくと分かりやすい.

1.1.4 分散 S^2 , 標準偏差 S

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

1.1.5 分散のより簡単な求め方

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

“「2乗の平均」-「平均の2乗」と覚えるべし.

1.1.6 単位の変換

$y_i = ax_i + b$ とすると,

$$\bar{y} = a\bar{x} + b$$

$$S_y^2 = a^2 S_x^2, \quad S_y = a S_x$$

1.1.7 基準化変量 z_i

$$z_i = \frac{x_i - \bar{x}}{S}$$

基準化変量は、「データが平均からどれだけずれているか」、さらに言えば、「標準偏差何個分ずれているか」を表す。
 z_i の平均は 0、分散は 1 である。

1.1.8 歪度 b_1 , 尖度 b_2

$$b_1 = \frac{1}{n} \sum_{i=1}^n z_i^3, \quad b_2 = \frac{1}{n} \sum_{i=1}^n z_i^4$$

b_1 は 0 と、 b_2 は 3 と大小を比較して使う。

1.1.9 変動係数 CV

$$CV = \frac{S}{\bar{x}}$$

異なるデータの散らばりを比較する際、標準偏差ではなくこの変動係数を使うほうが適切な場合がある。
(単位が異なるとき、平均の値が大きく異なるとき、など)

1.2 2次元データ

1.2.1 共分散 S_{xy}

$$S_{xy} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

1.2.2 共分散のより簡単な求め方

$$S_{xy} = \frac{1}{n} \sum_{i=1}^n x_i y_i - \bar{x} \bar{y}$$

“「積の平均」 - 「平均の積」” と覚えるべし。

1.2.3 相関係数 r_{xy}

$$r_{xy} = \frac{S_{xy}}{S_x S_y}$$

$0 < r_{xy} \leq 1 \iff$ 正の相関
 $r_{xy} \equiv 0 \iff$ 無相関
 $-1 \leq r_{xy} < 0 \iff$ 負の相関
 r_{xy} が ± 1 に近いほど相関が強い。

1.2.4 回帰直線

回帰直線 $y = a + bx$ は,

$$b = \frac{S_{xy}}{S_x^2}, \quad a = \bar{y} - b\bar{x}$$

で与えられる.

回帰直線って何ぞ, という人は教科書 P.77~ 82 を読んでください.

2 確率

2.1 基本公式

1. $P(A^c) = 1 - P(A)$
2. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

A^c は A の余事象. 2 つ目の式はベン図を描くと分かりやすい.

2.2 条件付確率

2.2.1 定義

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

↑は「 B が起こったという条件の下で A が起こる確率」である. 逆にしないよう注意すべし.

2.2.2 乗法公式

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

上の公式を変形しただけだが, $P(A \cap B)$ を求めたい場合はこちらが有用. 教科書 P.109 の例を参照のこと.

2.2.3 全確率公式

標本空間 Ω が

$$\Omega = H_1 \cup H_2 \cup \dots \cup H_n \quad (H_i \text{ は全て互いに背反})$$

と書けるとする.

このとき, 任意の事象 A の確率は,

$$P(A) = \sum_{i=1}^n P(A|H_i)P(H_i)$$

で与えられる.

教科書 P.110 の図が分かりやすい.

2.2.4 ベイズの定理

$$P(H_k|A) = \frac{P(A|H_k)P(H_k)}{\sum_{i=1}^n P(A|H_i)P(H_i)}$$

この公式は、 $P(A|H_k)$ などが分かっているときに $P(H_k|A)$ を求めたいとき、

例えば、「A 型の人が几帳面である確率」などが分かっているときに「几帳面な人が A 型である確率」を求めたいときに有用である。

ややこしいので、教科書 P.111~ 113 などをよく読み、十分理解しておくべきであろう。

2.3 事象の独立性

2.3.1 2個の事象の独立性

事象 A と B が独立 $\iff P(A \cap B) = P(A)P(B)$

さらに、これは $P(B) > 0$ のとき $P(A|B) = P(A)$ と同値で、

$P(B) > 0, P(B^c) > 0$ のとき $P(A|B) = P(A|B^c)$ と同値である。

2.3.2 n個の事象の独立性

事象 A_1, A_2, \dots, A_n が独立

$\iff A_1, A_2, \dots, A_n$ から、どのように任意個 (B_1, \dots, B_m) 選んでも、

$$P(B_1 \cap \dots \cap B_m) = P(B_1) \cdots P(B_m)$$

2.4 確率変数

2.4.1 離散型確率変数の平均 μ

$$\mu = E(X) = \sum_{k=1}^N x_k P(X = x_k)$$

2.4.2 X の関数 $g(X)$ の期待値 $E[g(X)]$

$$E[g(X)] = \sum_{k=1}^N g(x_k) P(X = x_k)$$

「期待値」と「平均」はほぼ同じようなものである。

2.4.3 平均は最小 2 乗値

$E[(X - c)^2]$ は $c = E(X)$ のとき最小。

2.4.4 離散型確率変数の分散 σ^2 と標準偏差 σ

$$\sigma^2 = V(X) = \sum_{k=1}^N (x_k - \mu)^2 P(X = x_k)$$

$$\sigma = D(X) = \sqrt{V(X)}$$

期待値の定義より, $V(X) = E[(x_k - \mu)^2]$ である.

2.4.5 分散のより簡単な求め方

$$V(X) = E(X^2) - \mu^2$$

やはり「2乗の平均」-「平均の2乗」である.

2.4.6 単位の変換

$$E(aX + b) = aE(X) + b \quad (\text{期待値の線形性})$$

$$V(aX + b) = a^2V(X), \quad D(aX + b) = aD(X)$$

2.4.7 基準化変量 Z

$$Z = \frac{X - \mu}{\sigma}$$

$E(Z) = 0, V(Z) = 1$ である.

基準化変量は, 主に正規分布を扱う上で非常に重要である.

2.4.8 連続型確率変数

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

$f(x)$ は X の確率密度関数である.

また, $F(a) = P(X \leq a) = \int_{-\infty}^a f(x)dx$ を X の分布関数という.

2.4.9 連続型確率変数の性質

$$P(X = a) = 0$$

例えば, 「身長が 170cm の人」がいたとしても,

より下の位まで見れば, その人の身長は 170.2cm かもしれないし, 170.002cm かもしれない.

そうやって突き詰めて考えると, 結局「身長がぴったり 170cm の人」は本質的には存在しないのである.

2.4.10 連続型確率変数の平均, 分散, 標準偏差

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx, \quad E[g(x)] = \int_{-\infty}^{\infty} g(x)f(x)dx,$$

$$V(X) = \int_{-\infty}^{\infty} (x - \mu)^2 f(x)dx, \quad D(X) = \sqrt{V(X)}$$

3 確率分布

3.1 色々な確率分布

3.1.1 離散型確率分布

名前	記号	$P(X = x)$	平均 $E(X)$	分散 $V(X)$	標準偏差 $D(X)$
ベルヌーイ分布	$Ber(p)$	$p^x(1-p)^{1-x}$	p	$p(1-p)$	$\sqrt{p(1-p)}$
二項分布	$B(n, p)$	${}_nC_x p^x(1-p)^{n-x}$	np	$np(1-p)$	$\sqrt{np(1-p)}$
ポアソン分布	$Po(\lambda)$	$e^{-\lambda} \frac{\lambda^x}{x!}$	λ	λ	$\sqrt{\lambda}$
幾何分布	$Ge(p)$	$p(1-p)^{x-1}$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$\frac{\sqrt{1-p}}{p}$

ベルヌーイ分布は二項分布で $n=1$ としたもの。

ポアソン分布は二項分布で $n \rightarrow \infty, p \rightarrow 0$ としたもの。したがって n が大きく p が小さいときに有用。

3.1.2 連続型確率分布

名前	記号	密度関数 $f(x)$	平均 $E(X)$	分散 $V(X)$	標準偏差 $D(X)$
一様分布	$U(a, b)$	$\begin{cases} \frac{1}{b-a} & (a \leq x \leq b) \\ 0 & (\text{その他}) \end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{b-a}{2\sqrt{3}}$
正規分布	$N(\mu, \sigma^2)$	$\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2	σ
指数分布	$Ex(\lambda)$	$\begin{cases} \lambda e^{-\lambda x} & (x > 0) \\ 0 & (x \leq 0) \end{cases}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\frac{1}{\lambda}$

他、後述するカイ 2 乗分布や t 分布などがある。

3.2 諸公式

3.2.1 幾何分布, 指数分布の無記憶性

幾何分布は次の式を満たす。

$$P(X = a + b | X > b) = P(X = a)$$

また, 指数分布は次の式を満たす。

$$P(X > a + b | X > b) = P(X > a)$$

無記憶性の例:

店員 A 「客来ないなあ」

店員 B 「もう 2 時間経つのに 1 人も来ないとか…」

店員 A 「いい加減来てもいい頃じゃね?」

店員 C 「それは違うね! 指数分布の無記憶性により, 君たちが 2 時間待ったかどうかに関わらず, 客の来る確率は変わらないのさ!」

店員 A 「…」

店員 B 「…」

無記憶性は幾何分布と指数分布に特有の性質である。

3.2.2 1次変換と正規分布

$X \sim N(\mu, \sigma^2)$ ならば,

$$aX + b \sim N(a\mu + b, a^2\sigma^2)$$

重要なのは、「 $aX + b$ も正規分布に従う」ということである。

3.2.3 正規分布の確率計算

※必ずテストに出るところです！

標準正規分布 $N(0, 1)$ の分布関数を Φ で表す。

$X \sim N(\mu, \sigma^2)$ ならば,

$$P(a \leq X \leq b) = \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right)$$

この公式は暗記するだけでなく、実際に使えるように計算練習を積んでおくべきである。

しつこいようだが、必ずテストに出るところである。

なお、正規分布表には $a < 0$ の場合の $\Phi(a)$ が載っていないが、

$\Phi(-a) = 1 - \Phi(a)$ である。これも覚えておくべし。

4 独立同一分布

4.1 同時確率分布

4.1.1 周辺分布

$$P(X = x_i) = \sum_{j=1}^N P(X = x_i, Y = y_j)$$

これを $f_X(x_i)$ と表し、 X の周辺分布という。

4.1.2 期待値

$$E[g(X, Y)] = \sum_{i=1}^M \sum_{j=1}^N g(x_i, y_j) P(X = x_i, Y = y_j)$$

X や Y のみに注目して、例えば $g(X, Y) = X$ とおくことも可能。→次

4.1.3 平均, 分散

$$\mu_X = E(X) = \sum_{i=1}^M \sum_{j=1}^N x_i P(X = x_i, Y = y_j) = \sum_{i=1}^M x_i f_X(x_i)$$

$$\sigma_X^2 = V(X) = E[(X - \mu_X)^2]$$

Y についても同様。

4.1.4 共分散 σ_{XY} , 相関係数 ρ_{XY}

$$\sigma_{XY} = C(X, Y) = E[(X - \mu_X)(Y - \mu_Y)],$$

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

4.2 独立性

4.2.1 独立性 (2次元の場合)

$X \perp\!\!\!\perp Y$ (X と Y が独立)

\iff 全ての i, j に対して,

$$P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j)$$

4.2.2 独立性 (n 次元の場合)

$X_1, X_2, \dots, X_n : \perp\!\!\!\perp$ (X_1, X_2, \dots, X_n が独立)

\iff 全ての (a_1, a_2, \dots, a_n) に対して,

$$P(X_1 = a_1, X_2 = a_2, \dots, X_n = a_n) = P(X_1 = a_1)P(X_2 = a_2) \cdots P(X_n = a_n)$$

4.2.3 独立性に関する定理

1. $X \perp\!\!\!\perp Y \Rightarrow E(XY) = E(X)E(Y)$
2. $X_1, X_2, \dots, X_n : \perp\!\!\!\perp \Rightarrow E(X_1 X_2 \cdots X_n) = E(X_1)E(X_2) \cdots E(X_n)$
3. 独立 \Rightarrow 無相関. すなわち, $X \perp\!\!\!\perp Y \Rightarrow C(X, Y) = 0$

4.3 和の分布

4.3.1 和の平均と分散

$X_1, X_2, \dots, X_n : \perp\!\!\!\perp$ のとき,

$$E(X_1 + X_2 + \cdots + X_n) = E(X_1) + E(X_2) + \cdots + E(X_n),$$

$$V(X_1 + X_2 + \cdots + X_n) = V(X_1) + V(X_2) + \cdots + V(X_n)$$

特に 2次元のとき,

$$E(X + Y) = E(X) + E(Y), \quad V(X + Y) = V(X) + V(Y)$$

4.3.2 標本平均の平均と分散

X_1, X_2, \dots, X_n は互いに独立で, 同一の分布 F に従うとする.

F の平均を μ , 分散を σ^2 とする.

このとき, 標本平均 $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ と, 和 $T = \sum_{i=1}^n X_i$ に関して次が成り立つ.

$$1. E(T) = n\mu, \quad V(T) = n\sigma^2$$

$$2. E(\bar{X}) = \mu, \quad V(\bar{X}) = \frac{\sigma^2}{n}$$

4.3.3 再生性

X_1, X_2, \dots, X_n は互いに独立で、それぞれ正規分布 $N(\mu_1, \sigma_1^2), N(\mu_2, \sigma_2^2), \dots, N(\mu_n, \sigma_n^2)$ に従うとする。このとき、

$$\sum_{i=1}^n X_i \sim N\left(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2\right)$$

正規分布に従う確率変数の和は、再び正規分布に従う、ということ。

正規分布の他、ベルヌーイ分布, 2 項分布, ポアソン分布などが再生性を持つ。

4.3.4 標本平均の分布

X_1, X_2, \dots, X_n は互いに独立で、同一の正規分布 $N(\mu, \sigma^2)$ に従うとする。このとき、

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

4.3.5 中心極限定理

X_1, X_2, \dots, X_n は互いに独立で、同一の分布 F (平均 μ , 分散 σ^2) に従うとする。

ここで $n \rightarrow \infty$ とすると、

\bar{X} の分布は $N\left(\mu, \frac{\sigma^2}{n}\right)$ に限りなく近づく。

これを用いた確率計算が以前に出題されている。

教科書 P.188,189 はよんでおくべし。