暗記したくない人のためのアクチュアリー数学

最終更新: 2022 年 9 月 21 日

これは何か

試験に合格するためには，時間制限の都合上ある程度暗記に頼らなくてはなりません．私が暗記量を減らそうとした努力の中身を公開しておきます．

注意

- 突然 X とかが出てきたら X はその項の分布にしたがう確率変数です．
- 定性的な考察でショートカットしている部分は「っぽい」と書いています．
- X_i, X_j とか書いているものは互いに独立です．
- $0 \le p \le 1$ に対して $1-p$ のことを断りなく q と書きます．

離散一様分布 $DU\{1, 2, \cdots, n\}$

n 面さいころを 1 回ふって, 何の目が出るか？ がしたがう分布．

- 確率関数: n 択一様なので, $P_X(x) = 1/n$.
- 期待値: ど真ん中が期待値っぽいので, $(n+1)/2$ （中点）.
- 分散: 仕方なく覚える．$(n+1)(n-1)/12$.

連続一様分布 $U(a, b)$

- 密度関数: (a, b) で積分して 1 になるように, $f_X(x) = 1/(b-a)$.
- 分布関数: 積分.
- 期待値: ど真ん中が期待値っぽいので, $(a+b)/2$ （中点）.
- 分散: 仕方なく覚える．$(b-a)^2/12$. 離散の場合と 2 乗のかかり方が違うので注意.
- モーメント母関数: 定義通り計算する.

二項分布 $Bin(n, p)$

成功率 p の何かを n 回やって, 何回成功するか？ がしたがう分布．$n = 1$ のときの分布をとくに **Bernoulli 分布**といって, $Ber(p)$ とかく．$Ber(p)$ の特性値を出してからそれをいじるのがよい.

- $Ber(p)$ の特性値（計算する）
  - 期待値: $p \cdot 1 + q \cdot 0 = p$.
  - 分散: $E[X^2] - E[X]^2 = p - p^2 = pq$. $p = 0, 1$ のとき 0 になることに注意.
  - モーメント母関数: $pe^t + qe^0 = pe^t + q$.

以上より, 二項分布について

- 確率関数: 数 A.

- 再生性: $X_1 \sim Ber(n_1, p), X_2 \sim Ber(n_2, p)$ ならば $X_1 + X_2 \sim Ber(n_1 + n_2, p)$. 今後, このような意味で n に関する再生性などと言う.
- 期待値: n 回やったらだいたい np 回成功しそうなので.
- 分散: 再生性と分散の性質 $V(X_1 + X_2) = V(X_1) + V(X_2)$ と Bernoulli 分布の分散より.
- モーメント母関数: 再生性より, Bernoulli 分布の母関数の n 乗.

Poisson 分布 $Po(\lambda)$

二項分布において平均 $np = \lambda$ を固定して $n \to \infty$, $p \to 0$ としたもの. 成功率 p がきわめて低い試行を大量に行うときに出てくる. 2 つの極限を同時にとるため成功率と試行回数のバランスがとれて, 平均の成功回数 λ は有限になるというわけ.

- 確率関数: e^{λ} の Taylor 展開を書いてから, $e^{-\lambda}$ をかけて規格化する.
- 再生性: 二項分布の極限なので λ に関する再生性として引き継がれる.
- 期待値: 定義より.
- 分散: $npq \to np = \lambda$. 同様に, キュムラントに関する一般的な事実 $C_X^{(n)}(0) = \lambda$ も成り立つ.
- モーメント母関数: 定義通り計算する. $(e^t\lambda)^x$ の形を作って e^x の Taylor 展開を用いる.

負の二項分布 $NB(\alpha, p)$

成功率 p の何かが α 回成功するまでに何回失敗したか? がしたがう分布. 謎のネーミングだが, 負の二項定理と関係していることからこの名前がついている. $\alpha = 1$ のときの分布をとくに **幾何分布** といって, $Geo(p)$ とかく. $Geo(p)$ の特性値を出してからそれをいじるのがよい.

- $Geo(p)$ の特性値
  - 確率関数: 数 A.
  - 期待値: 試行回数の期待値は $1/p$ っぽいから（成功確率 $1/5$ の何かを 5 回やると 1 回くらいは成功しそう）, 最後の成功ターンを引いて $1 - 1/p = q/p$.
  - 分散: 覚える q/p^2. 忘れたらモーメント母関数から計算. $p = 1$ のときは必ず $X = 0$ であるから分散は 0 になることを確認.
  - モーメント母関数: 定義通り計算して 10 秒.

以上より, 負の二項分布について

- 確率関数: 数 A.
- 再生性: α に関する再生性をもつ. 1 回成功するたびに失敗回数がリセットすると思えばよい.
- 期待値: 幾何分布の期待値 + 再生性.
- 分散: 幾何分布の分散 + 再生性.
- モーメント母関数: 再生性より, 幾何分布の母関数の α 乗.

超幾何分布 $H(N, m, n)$

N 個（赤 m 個, 白 $N - m$ 個）のボールが入った箱から無作為に n 個抽出したときにいくつ赤玉が含まれるか? がしたがう分布. **非復元抽出** である点が二項分布と異なる. 二項分布との対応としては, $p \iff m/N$.

- 確率関数: 数 A.
- 期待値: 二項分布と同じ．直感的じゃないと思う.
- 分散: 二項分布の分散 $N \cdot (m/N)(1-m/N)$ に有限母集団修正係数 $(N-n)/(N-1)$ をかける．この係数は $O(1)$ のはずで（$N \to \infty$ で二項分布の分散に収束しなければならないから），$N=n$ でゼロになる（ボールを全部とったら必ず m 個の赤玉が含まれるから）ことから決まる．分母が N ではなく $N-1$ なのは覚える.

正規分布 $\mathcal{N}(\mu, \sigma^2)$

- 密度関数: 理系大学生は勝手に覚えている.
- 再生性: μ, σ^2 両方について再生性をもつ.
- モーメント母関数: たたくとモーメントが落ちてくるように係数を決める.

対数正規分布 $LN(\mu, \sigma^2)$

- 密度関数: 覚えてもいいけど, 対数正規分布の定義 $X \sim \mathcal{N}(\mu, \sigma^2)$ のとき $e^Y \sim LN(\mu, \sigma^2)$ から変数変換してもそんなに時間かからない．この定義がそのまま問題になっていることが多い.
- 平均, 分散: 上の定義を念頭に正規分布のモーメント母関数をいじる（$M_X(t)$ の t に 1 とか 2 を代入する）.

ガンマ分布 $\Gamma(\alpha, \beta)$

期間 $1/\beta$ ごとに 1 回程度起こるランダムな事象が α 回おこるまでにかかる時間がしたがう分布．α は非負整数に限らないので, 気持ちとしては負の二項分布の連続バージョンである．いろんなものがガンマ分布やガンマ関数に帰着できて便利, これを自然に使えるようになると覚えることが減ってよい．$\alpha = 1$ のときの分布をとくに **指数分布** といって, $Ex(\beta)$ とかく．指数分布の特性値を出してからそれをいじるのがよい.

- $Ex(\beta)$ の特性値
  - 確率関数: $f_X(x) = \beta e^{-\beta x}$
  - 期待値: 分布の意味より.
  - 分散: 覚える, $1/\beta^2$．幾何分布の連続バージョンで, 似てる.
  - モーメント母関数: 定義通り計算.

以上より, ガンマ分布について

- 密度関数: ガンマ関数の定義から規格化する．ガンマ関数の定義はがんばって覚える.
- 再生性: α に関する再生性.
- 期待値: 指数分布の期待値 + 再生性.
- 分散: 指数分布の分散 + 再生性.
- モーメント母関数: 再生性より, 指数分布の母関数の α 乗．しかしモーメントがほしいときはモーメントを直接求めに行ったほうがはやい.
- キュムラント: 覚えるかなあ...

とくに $\Gamma(n/2, 1/2)$ を自由度 n の **カイ 2 乗分布** といって $\chi^2(n)$ とかく．これは大事で, 標準正規分布にしたがう n 個の独立な確率変数の 2 乗和 $\sim \chi^2(n)$.

ベータ分布 $Beta(p, q)$

表が出る確率がわからないコイン投げにおける実験結果が 表 $p-1$ 回, 裏 $q-1$ 回だったときの, 表が出る確率 x の予測分布. ベータ分布の意味付けについては [1] がわかりやすい.

- 密度関数: ベータ関数の定義から規格化する. ベータ関数の定義はがんばって覚える.
- 期待値: ベータ分布の意味から.
- 分散: 下でモーメントを求めてから計算する.
- モーメント: 直接計算するとガンマ関数でかける.
- $1-X \sim Beta(q, p)$: 表裏の回数が逆になると確率の予想も逆になるので.

パレート分布 $Pa(\alpha, \beta)$

- 密度関数: $\alpha\beta^\alpha/x^{\alpha+1}$ 分布関数が覚えやすいので, そっちを覚えて微分する.
- 分布関数: $1-(\beta/x)^\alpha$. 覚えやすい. $x \to \infty$ で 1 に行き, $\beta > x$ で負になるので定義域は $\beta \leq x$.
- 期待値: 定義から計算. $\alpha > 1$ で存在.
- 分散: 定義から計算. $\alpha > 2$ で存在.
- 定義域が 0 からはじまるように平行移動したものを第 2 種パレート分布 $Pa2(\alpha, \beta)$ という. 分布関数 $1-(\beta/(x+\beta))^\alpha$

t 分布 $t(n)$

<u>重要な性質</u>: $X \sim \mathcal{N}(0,1), Y \sim \chi^2(n)$ のとき $\frac{X}{\sqrt{Y/n}} \sim t(n)$. $n=1$ で Cauchy 分布, $n \to \infty$ で $\mathcal{N}(0,1)$ に分布収束（$n \to \infty$ の性質）.

- 密度関数: $\frac{1}{\sqrt{n}B(n/2,1/2)}(1+\frac{x^2}{n})^{-\frac{n+1}{2}}$. 覚える. 後半部分は, $n \to \infty$ の性質より, 極限で $e^{-x^2/2}$ に行くように書く. ベータ関数の引数が $\chi^2(n)$ のガンマ分布表示の引数と同じ.
- 期待値: 0. 「重要な性質」からあきらか.
- 分散: $n/(n-2)$. $n \to \infty$ の性質より, この極限で 1 に行くようにする.

F 分布 $F(m, n)$

<u>重要な性質</u>: $X \sim \chi^2(m), Y \sim \chi^2(n)$ のとき $\frac{X/m}{Y/n} \sim F(m,n)$. これよりただちに $X \sim F(m,n) \iff 1/X \sim F(n,m)$.

<u>実用上重要な性質</u>: $F_n^m(\alpha) = 1/F_m^n(1-\alpha)$. 分布の表には片方しか載ってないことがあるから, これを使って変換する.

- 密度関数: $\frac{1}{B(m/2,n/2)}(\frac{m}{n})^{m/2}x^{m/2-1}(1+\frac{m}{n}x)^{-(m+n)/2}$. 覚える. つねに m が先.
- 期待値: 覚える. m によらない.
- 分散: どうすれば...

コーシー分布 $C(\mu, \sigma)$

$X, Y \sim \mathcal{N}(0,1)$ のとき, $X/Y \sim C(0,1)$.

- 密度関数: $\frac{1}{\pi}\frac{\sigma}{\sigma^2+(x-\mu)^2}$.
- 期待値, 分散: 可積分でない.

参考文献

[1] 高校数学の美しい物語, `https://manabitimes.jp/math/1267`

[2] 久保川達也, 現代数理統計学の基礎, 共立出版.

[3] 竹村彰通, 現代数理統計学, 学術図書出版社.