チェビシェフの不等式と大数の (弱) 法則

樋口さぶろお

龍谷大学理工学部数理情報学科

確率統計☆演習 II L05(2015-05-08 Fri)

最終更新: Time-stamp: "2015-05-08 Fri 08:16 JST hig"

今日の目標

- 離散確率変数についてのベイズ推定ができる
- 大数の (弱) 法則の正確な表現と意味が説明できる

http://hig3.net

L04-S1
Quiz 解答:ピアソンの χ^2 と独立性の検定

1. 期待度数は

	右利き	右利きでない	計
早生まれ	$\frac{4}{3}$	$\frac{2}{3}$	2
早生まれでない	$\frac{8}{3}$	$\frac{4}{3}$	4
計	4	2	6

$$\chi^2 = \frac{(1-\frac{4}{3})^2}{\frac{4}{3}} + \frac{(1-\frac{2}{3})^2}{\frac{2}{3}} + \frac{(3-\frac{8}{3})^2}{\frac{8}{3}} + \frac{(1-\frac{4}{3})^2}{\frac{4}{3}} = \frac{3}{8}.$$

2. 自由度は $k=(r-1)\times(c-1)=1$. 有意水準 $\alpha=0.05$ で, $\chi_\alpha(k)=3.841>\frac{3}{8}$. よって, 独立であるという帰無仮説は棄却できない.

チェビシェフの不等式

チェビシェフの不等式 Chebyshev's inequality

X: 離散型または連続型確率変数
$\mu = \mathrm{E}[X]$: 母平均値
$\sigma^2 = \mathrm{V}[X]$: 母分散
$a > 0$: 任意の正の実数
のとき次が成立する.

$$P(|X - \mu| \geq a\sigma) \leq \frac{1}{a^2}$$

どんな X にも使えて便利な不等式.
意味

チェビシェフの不等式の証明

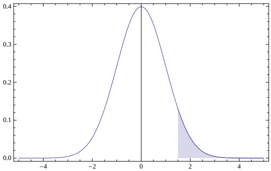
L05-Q1

Quiz(チェビシェフの不等式の正規分布への適用例)

確率変数 X が標準正規分布 $\mathrm{N}(0,1^2)$ にしたがうとき, チェビシェフの不等式の両辺の値を, $a=1,2,3$ について計算してみよう.

標準正規確率表 (上側確率=$Q(x) = 1 - \Phi(x)$)

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010



ここまで来たよ

1. 略解: カテゴリ変数と独立性の検定
2. チェビシェフの不等式と大数の (弱) 法則
   - 大数の (弱) 法則
3. ベイズ推定
   - ベイズの公式

大数の (弱) 法則

大数の (弱) 法則 law or large numbers

$X_1, \ldots, X_n$: 独立同分布に従う確率変数 (‘サイズ n の標本’) $\mu = \mathrm{E}[X_i]$
確率変数 $\overline{X}_{(n)} = \frac{1}{n}(X_1 + \cdots + X_n)$ (‘標本平均値’)
$\epsilon > 0$: 任意の (どんなに小さくてもよい) 正の数
のとき,

$$P(|\overline{X}_{(n)} - \mu| > \epsilon) \to 0 \quad (n \to +\infty)$$

$\overline{X}_{(n)}$ は μ に確率収束する という.

普通の「a_n が a に収束する」とは　集合位相

$$\forall \epsilon_1 > 0 \quad \exists n_1 \quad \text{s.t.} \quad n > n_1 \Rightarrow |a_n - a| < \epsilon_1$$

ϵ と ϵ_1 は別物. 全部きちんと書くと,

$$\forall \epsilon > 0 \quad \forall \epsilon_1 > 0 \quad \exists n_1 \quad \text{s.t.} \quad n > n_1 \Rightarrow |P(|\overline{X}_{(n)} - \mu| > \epsilon) - 0| < \epsilon_1$$

大数の (強) 法則 もある 「…の確率がぴったり 0」強収束, 概収束

大数の (弱) 法則の意味

$X, Y = \phi(X)$, $Z = \mathbf{1}_{[条件]}(X)$
などに適用すると, 平均値や期待値や比率 (確率) についても,
標本サイズ $n \to +\infty$ で, ‘標本期待値’ $\frac{1}{n}(X_1 + \cdots + X_n)$ による母期待値の推定が正確になることがわかる.

大数の (弱) 法則の証明

ここまで来たよ

1. 略解: カテゴリ変数と独立性の検定
2. チェビシェフの不等式と大数の (弱) 法則
   - 大数の (弱) 法則
3. ベイズ推定
   - ベイズの公式

条件付き確率の計算 I

L05-Q2

Quiz(ベイズの公式)

抽選用の袋に何個かの色つきボールが入っており, ボールを割ると, $1/4$ に当たりの紙が入っている
当たりのボールのうち赤いボールが $1/3$, 白いボールが $2/3$ である.
外れのボールのうち赤いボールが $4/9$, 白いボールが $5/9$ である.
無作為に袋から取り出したボールが赤い場合, 当たりである確率を求めよう.

ベイズの公式

$$P(X=x|Y=y)=\frac{P(Y=y|X=x)P(X=x)}{\sum_x P(Y=y|X=x)P(X=x)}. \tag{5.1}$$

$$P(Y=y|X=x)=\frac{P(X=x|Y=y)P(Y=y)}{\sum_y P(X=x|Y=y)P(Y=y)}. \tag{5.2}$$

$P(X=x|Y=y)$ を $P(Y=y|X=x)$ で書き表す式, およびその逆の式.

L05-Q3

Quiz(ベイズの公式)

外見で区別できない, 品種 1(甘い) と品種 2(渋い) の柿がかごに入っている.
品種 1 は, 確率 0.95 で赤に, 確率 0.05 で黄色になる.
品種 2 は, 確率 0.125 で赤に, 確率 0.875 で黄色になる.
確率変数 X, Y を用いて, 品種 1(甘い) を $X = 1$, 品種 2(渋い) を $X = 2$, 赤いを $Y = 10$, 黄色いを $Y = 20$ と表現する.

1. 問題文から $P(Y = y|X = x)$ を読み取ろう.
2. かごの柿の $1/5$ が甘い柿であるとする. いま, 無作為に 1 個の柿を取りだしたところ, 赤い柿だった. ベイズの公式を使って, 取り出した赤い柿が甘い確率 $P(X = 1|Y = 10)$ を求めよう.
3. 仮にかごの柿の $1/5$ が渋い柿であるとする. いま, 無作為に 1 個の柿を取りだしたところ, 黄色い柿だった. ベイズの公式を使って, 取り出した黄色い柿が渋い確率を求めよう.

ベイズ的な考え方

$\boxed{\text{事後確率 } P(X=x|Y=y)}$ $\longleftarrow$ $\boxed{\text{事前確率 } P(X=x)}$

$\uparrow$

$\boxed{\text{情報 } Y=y}$

主観確率

ベイズの定理=ベイズの公式 (+ニュアンス?)

L05-Q4

Quiz(ベイズの推定)

外見で区別できない, 品種 1(甘い) と品種 2(渋い) の柿がかごに入っている.
品種 1 は, 確率 0.95 で赤に, 確率 0.05 で黄色になる.
品種 2 は, 確率 0.125 で赤に, 確率 0.875 で黄色になる.
確率変数 X, Y を用いて, 品種 1(甘い) を $X = 1$, 品種 2(渋い) を $X = 2$, 赤いを $Y = 10$, 黄色いを $Y = 20$ と表現する.

1. いろんな情報と信念から, かごの柿の $1/5$ が甘い柿だろうと思っている. したがって, 無作為に取り出した柿 (色を見ていない) が甘い確率 (事前確率) は $1/5$ である. 色を見たところ, 赤い柿だった. この情報を得た後では, その柿が甘い確率 (事後確率) は?
2. いろんな情報と信念から, かごの柿の $1/5$ が渋い柿だろうと思っている. したがって, 無作為に取り出した柿 (色を見ていない) が渋い確率 (事前確率) は $1/5$ である. 色を見たところ, 黄色い柿だった. この情報を得た後では, その柿が渋い確率 (事後確率) は?

連絡

確率統計☆演習 II プチテスト 2015-05-29 金 3
数理情報学科オープンレクチャー 2015-05-20 水 11:05-12:35 1 号館 1-534
台湾への留学と中枢神経における呼吸モデルについて
Math ラウンジ=チューター月火水木昼, 1-612 1-614
統計検定ウィーク 2015-05-11 – 2015-05-14. 過去問や問題集お見せします. 受検の相談にのります.
その他の質問・相談にもふだん通り歓迎です.
統計検定を取ろう! 2015-05-15 申込締切, 2015-06-21 検定実施
2 級 or 3 級をお奨めします `http://www.toukei-kentei.jp/`

manaba 出席カード提出
`https://attend.ryukoku.ac.jp`