尤度比検定

樋口さぶろお

龍谷大学大学院理工学研究科数理情報学専攻

理論物理学特論 L04(2015-10-15 Thu)

最終更新: Time-stamp: "2015-10-16 Fri 18:24 JST hig"

今日の目標

1. 検定のネイマン-ピアソンの枠組を説明できる
2. (自分でプログラムして) パラメトリックブートストラップ法による検定を実行できる
3. カイ二乗分布と正規分布の関係を説明できる
4. 数表からカイ二乗分布の確率を求められる

http://hig3.net

L03-Q1

Quiz 解答:ポアソン回帰

$$
\begin{aligned}
\log L =& \log \frac{\mathrm{e}^{(\beta_1+\beta_2+\beta_3)\cdot 1}}{1!}\mathrm{e}^{-\mathrm{e}^{\beta_1+\beta_2+\beta_3}} + \log \frac{\mathrm{e}^{(\beta_1+2\beta_2)\cdot 3}}{3!}\mathrm{e}^{-\mathrm{e}^{\beta_1+2\beta_2}} \\
&+ \log \frac{\mathrm{e}^{(\beta_1+2\beta_2)\cdot 5}}{5!}\mathrm{e}^{-\mathrm{e}^{\beta_1+2\beta_2}} + \log \frac{\mathrm{e}^{(\beta_1+3\beta_2+\beta_3)\cdot 8}}{8!}\mathrm{e}^{-\mathrm{e}^{\beta_1+3\beta_2+\beta_3}} \\
=&(\beta_1+\beta_2+\beta_3)\cdot 1 - \mathrm{e}^{\beta_1+\beta_2+\beta_3} + (\beta_1+2\beta_2)\cdot 3 - \mathrm{e}^{\beta_1+2\beta_2} \\
&+ (\beta_1+2\beta_2)\cdot 5 - \mathrm{e}^{\beta_1+2\beta_2} + (\beta_1+3\beta_2+\beta_3)\cdot 8 - \mathrm{e}^{\beta_1+3\beta_2+\beta_3}.
\end{aligned}
$$

ここまで来たよ

1. 略解:逸脱度・AIC
2. 尤度比検定
   - カイ 2 乗分布

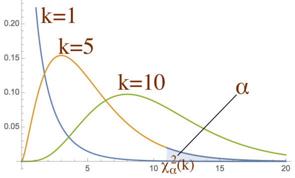
カイ 2 乗分布

つぎの確率密度関数をもつ連続型確率変数 X を, 自由度 n のカイ 2 乗分布 $\mathrm{Ga}(\frac{n}{2}, \frac{1}{2})$ にしたがうという.

$$f(x) = \begin{cases} \frac{1}{C_n} x^{\frac{n}{2}-1} \mathrm{e}^{-\frac{x}{2}} & (x > 0) \\ 0 & (\text{他}) \end{cases}$$

ただし, $C_n = \int_0^\infty x^{\frac{n}{2}-1} \mathrm{e}^{-\frac{x}{2}} \mathrm{d}x$.

意味:
あとで見るが, $Z_i \sim \mathrm{N}(0, 1^2)$ のとき, $T \sim Z_1^2 + \cdots + Z_n^2 \sim \mathrm{Ga}(\frac{n}{2}, \frac{1}{2})$.



$n = 1, 5, 10.$

カイ 2 乗分布のモーメント母関数と期待値

X が自由度 n のカイ 2 乗分布にしたがうとき,

$$M_X(\lambda) = (1 - 2\lambda)^{-\frac{n}{2}}$$

$$\mathrm{E}[X] = \boxed{\quad}, \qquad \mathrm{V}[X] = \boxed{\quad\quad}$$

正規分布とカイ 2 乗分布の関係

$Z \sim \mathrm{N}(0, 1^2)$ とする.

正規分布とカイ 2 乗分布の関係

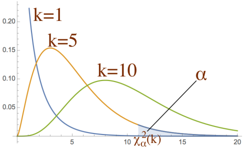
$Z_i \sim \mathrm{N}(0, 1^2)$ のとき, $T = Z_1^2 + \cdots + Z_n^2 \sim \mathrm{Ga}(\frac{n}{2}, \frac{1}{2})$.

上の T のことを χ^2 と書く記法も見かける.

χ^2 分布表

$\alpha = P(\chi^2 > \chi^2_\alpha(k))$.

$k\backslash\alpha$	0.995	0.99	0.975	0.95	0.9	0.1	0.05	0.025	0.01	0.005
1	0.00003927	0.0001571	0.0009821	0.003932	0.01579	2.706	3.841	5.024	6.635	7.879
2	0.01003	0.02010	0.05064	0.1026	0.2107	4.605	5.991	7.378	9.210	10.60
3	0.07172	0.1148	0.2158	0.3518	0.5844	6.251	7.815	9.348	11.34	12.84
4	0.2070	0.2971	0.4844	0.7107	1.064	7.779	9.488	11.14	13.28	14.86
5	0.4117	0.5543	0.8312	1.145	1.610	9.236	11.07	12.83	15.09	16.75
6	0.6757	0.8721	1.237	1.635	2.204	10.64	12.59	14.45	16.81	18.55
7	0.9893	1.239	1.690	2.167	2.833	12.02	14.07	16.01	18.48	20.28
8	1.344	1.646	2.180	2.733	3.490	13.36	15.51	17.53	20.09	21.95
9	1.735	2.088	2.700	3.325	4.168	14.68	16.92	19.02	21.67	23.59
10	2.156	2.558	3.247	3.940	4.865	15.99	18.31	20.48	23.21	25.19
11	2.603	3.053	3.816	4.575	5.578	17.28	19.68	21.92	24.72	26.76
12	3.074	3.571	4.404	5.226	6.304	18.55	21.03	23.34	26.22	28.30
13	3.565	4.107	5.009	5.892	7.042	19.81	22.36	24.74	27.69	29.82
14	4.075	4.660	5.629	6.571	7.790	21.06	23.68	26.12	29.14	31.32
15	4.601	5.229	6.262	7.261	8.547	22.31	25.00	27.49	30.58	32.80
16	5.142	5.812	6.908	7.962	9.312	23.54	26.30	28.85	32.00	34.27
17	5.697	6.408	7.564	8.672	10.09	24.77	27.59	30.19	33.41	35.72
18	6.265	7.015	8.231	9.390	10.86	25.99	28.87	31.53	34.81	37.16
19	6.844	7.633	8.907	10.12	11.65	27.20	30.14	32.85	36.19	38.58
20	7.434	8.260	9.591	10.85	12.44	28.41	31.41	34.17	37.57	40.00
30	13.79	14.95	16.79	18.49	20.60	40.26	43.77	46.98	50.89	53.67
40	20.71	22.16	24.43	26.51	29.05	51.81	55.76	59.34	63.69	66.77
50	27.99	29.71	32.36	34.76	37.69	63.17	67.50	71.42	76.15	79.49
60	35.53	37.48	40.48	43.19	46.46	74.40	79.08	83.30	88.38	91.95
70	43.28	45.44	48.76	51.74	55.33	85.53	90.53	95.02	100.4	104.2
80	51.17	53.54	57.15	60.39	64.28	96.58	101.9	106.6	112.3	116.3
90	59.20	61.75	65.65	69.13	73.29	107.6	113.1	118.1	124.1	128.3
100	67.33	70.06	74.22	77.93	82.36	118.5	124.3	129.6	135.8	140.2



Quiz(カイ 2 乗分布)

自由度 $n=1$ の カイ 2 乗分布にしたがう確率変数 T を考える.

1. $P(T>a)=0.05$ となる a を求めよう.
2. $P(T<b)=0.05$ となる a を求めよう.

を求めよう.

L04-Q2

Quiz(カイ 2 乗分布)

標準正規分布 $\mathrm{N}(0,1^2)$ にしたがう独立な確率変数 Z_1,Z_2,Z_3 を考える.

1. $\mathrm{E}[Z_1^2+Z_2^2+Z_3^2], \mathrm{V}[Z_1^2+Z_2^2+Z_3^2]$ を答えよう.
2. $P(Z_1^2+Z_2^2+Z_3^2>a)=0.05$ となる a の値を求めよう.
3. $P(Z_1^2+Z_2^2<b)=0.01$ となる b の値を求めよう.

を求めよう.

L04-Q3

Quiz(カイ 2 乗分布)

正規分布 $\mathrm{N}(0, 2^2)$ にしたがう独立な確率変数 X_1, X_2, X_3, X_4, X_5 を考える.

1. $\mathrm{E}[\frac{1}{5}[X_1^2 + \cdots + X_5^2]], \mathrm{V}[\frac{1}{5}[X_1^2 + \cdots + X_5^2]]$ を答えよう.
2. $P(\frac{1}{5}[X_1^2 + \cdots + X_5^2] > a) = 0.05$ となる a の値を求めよう.
3. $P(\frac{1}{5}[X_1^2 + \cdots + X_5^2] < b) = 0.01$ となる b の値を求めよう.

を求めよう.

正規分布とカイ 2 乗分布

母平均値 μ, 母分散 σ の正規分布にしたがう独立な X_i $(i = 1, \ldots, n)$ に対して,

$$n \times \frac{1}{n} \left[\left(\frac{X_1 - \mu}{\sigma} \right)^2 + \cdots + \left(\frac{X_n - \mu}{\sigma} \right)^2 \right]$$

は自由度 n のカイ 2 乗分布にしたがう.

正規分布とカイ 2 乗分布

母平均値 μ, 母分散 σ の正規分布にしたがう独立な X_i $(i = 1, \ldots, n)$ に対して,

$$\frac{n-1}{\sigma^2} \times 不偏標本分散 \frac{1}{n-1}\left[\left(X_1 - \overline{X}\right)^2 + \cdots + \left(X_n - \overline{X}\right)^2\right]$$

は自由度 $n-1$ のカイ 2 乗分布にしたがう.
ここで, $\overline{X}$ は標本平均値 $\frac{1}{n}[X_1 + \cdots + X_n]$.

$P(a < \frac{(n-1)S^2}{\sigma^2} < b) = P(\frac{(n-1)S^2}{b} < \sigma^2 < \frac{(n-1)S^2}{a}) = 1 - \alpha.$

母分散の区間推定

正規分布にしたがう確率変数 X のサイズ n の標本の不偏標本分散が S^2 であるとき, 母分散 $\sigma^2 = \mathrm{V}[X]$ の信頼係数 $1 - \alpha$ の信頼区間は,

$$\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}$$

L04-Q4

Quiz(母分散の区間推定)

あるファーストフードチェーンのポテトフライ S の重さは正規分布に従うという.
お店で 9 個のポテトフライ S サイズを買って重さを量ったところ, 下のようだった (単位は g).

$$78, 78, 78, 78, 80, 82, 82, 82, 82$$

母平均値と母分散を信頼係数 $1 - \alpha = 0.95$ で区間推定しよう.

L04-Q5

Quiz(母分散の検定)

あるファーストフードチェーンのポテトフライ S の重さは, 母分散 $\sigma_0^2 = 4\mathrm{g}^2$ の分布であることが定められているという.
トレーニング中のアルバイトの人に, ポテトフライ S サイズを 9 個作ってもらったところ, 重さは下のようだった (単位は g).

$$76, 76, 76, 76, 80, 84, 84, 84, 84.$$

このアルバイトの作るポテトフライ S の重さの母分散 σ_1^2 は, σ_0^2 と異なるか? アルバイトのほうの重さが正規分布にしたがうと仮定し, 有意水準 5% で, 母分散の χ^2 検定で判定しよう.

連絡

- オフィスアワー月 4 木 6(1-502)