指数分布・カイ 2 乗分布

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　L11(2015-06-26 Fri)

最終更新: Time-stamp: "2015-06-26 Fri 18:55 JST hig"

今日の目標

- 指数分布にしたがう確率変数について, 確率, 母期待値が計算できる
- カイ二乗分布にしたがう確率変数について, 確率, 母期待値が計算できる

hig3.net

L10-Q5

Quiz 解答:指数分布

1. $\alpha = 0.3/$時間. $\mathrm{E}[X] = 1/0.3$ 時間$=\frac{10}{3}$ 時間, $\mathrm{V}[X] = 1/0.3^2$ 時間$^2 = \frac{100}{9}$ 時間^2.
2. $\int_2^\infty \alpha \mathrm{e}^{-\alpha x}\,\mathrm{d}x = [-\mathrm{e}^{-\alpha x}]_2^\infty = \mathrm{e}^{-0.3\times 2} = \mathrm{e}^{-0.6} = 0.549.$

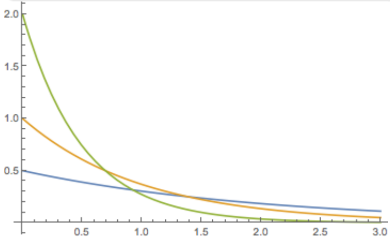
ここまで来たよ

1. 略解: 統計的仮説検定・指数分布
2. 指数分布・カイ 2 乗分布
   - 指数分布
   - カイ 2 乗分布

指数分布

つぎの確率密度関数をもつ連続型確率変数 X を, パラメタ $\alpha > 0$ の指数分布にしたがう, という.

$$f(x) = \begin{cases} \alpha \mathrm{e}^{-\alpha x} & (x > 0) \\ 0 & (\text{他}) \end{cases}$$



$\alpha = 0.5, 1, 2.$

意味: 独立で, 頻度が時間の長さに比例して (平均 α 回/分) 起きるできごと (その回数はポアソン分布にしたがう) の, おきる時間間隔 x 分の分布. 分のかわりに時, 秒, ゲームでも可.

指数分布のモーメント母関数と期待値

X がパラメタ α の指数分布にしたがうとき,

$$M_X(\lambda) = \frac{\alpha}{\alpha - \lambda} \qquad (\lambda < \alpha)$$

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

分布の間の関係

時間	回数	間隔
離散	2 項分布 (離散)	幾何分布 (離散)
	$\downarrow np = \alpha, n \to \infty$	
連続	ポアソン分布 (離散)	指数分布 (連続)

本来なら, 指数分布で考えるべきことを, 離散化して, 幾何分布で考えてみる.

時間 1 を n 個に等分する. 等分された時間帯に起きる確率 $p = \frac{\alpha}{n}$.

時刻 x つまり $k = nx$ 番目の区間に初めて発生する確率は,

$\frac{\alpha}{n}(1 - \frac{\alpha}{n})^{nx}$.

確率密度は, 区間の長さ $1/n$ で割って,

$\frac{1}{1/n}\frac{\alpha}{n}(1 - \frac{\alpha}{n})^{nx} \overset{n\to+\infty}{\to} \alpha \mathrm{e}^{-\alpha x}$.

L11-Q1

Quiz(指数分布)

あるサッカーチームは, 1 ゲーム 90 分で平均 4.5 点得点できる.

1. 得点と得点の時間間隔の母平均値を求めよう.
2. 得点と得点の時間間隔が 5 分未満である確率を求めよう.
3. 得点と得点の時間間隔が 15 分以上 25 分未満になる確率を求めよう.

ここまで来たよ

1. 略解: 統計的仮説検定・指数分布
2. 指数分布・カイ 2 乗分布
   - 指数分布
   - カイ 2 乗分布

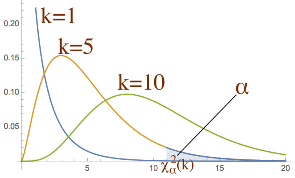
カイ 2 乗分布

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,

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

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

カイ 2 乗分布の再生性

$X_1 \sim \mathrm{Ga}(\frac{n_1}{2}, \frac{1}{2}), X_2 \sim \mathrm{Ga}(\frac{n_2}{2}, \frac{1}{2})$ のとき, $X_1 + X_2 \sim \mathrm{Ga}(\frac{n_1+n_2}{2}, \frac{1}{2})$

正規分布とカイ 2 乗分布の関係

$Z \sim \mathrm{N}(0, 1^2)$ とする.

$aZ + b \sim \mathrm{N}(\boxed{}, \boxed{})$ はすでに知っているが,

$Z^2 \sim \mathrm{Ga}(\frac{1}{2}, \frac{1}{2})$. なぜなら,

$$M_{Z^2}(\lambda) = \mathrm{E}[\mathrm{e}^{\lambda Z^2}] = \int_{-\infty}^{+\infty} \mathrm{e}^{\lambda z^2} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}z^2}\, \mathrm{d}z = \cdots = (1 - 2\lambda)^{-1/2}$$

正規分布とカイ 2 乗分布の関係

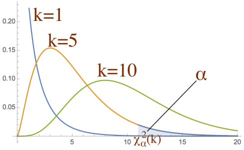
$Z_i \sim \mathrm{N}(0, 1^2)$ のとき, $T = Z_1^2 + \cdots + Z_n^2 \sim \mathrm{Ga}(\frac{n}{2}, \frac{1}{2})$.

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



L11-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.

L11-Q3

Quiz(カイ 2 乗分布)

標準正規分布 $\mathrm{N}(0,2^2)$ 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,

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

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 $\frac{1}{n}[X_1 + \cdots + X_n]$.

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

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,

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

L11-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.

L11-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.

Math ラウンジ=チューター
月火水木昼, 1-614
各科目のレポート, 課題などその他の質問・相談歓迎です.
スケジュール
2015-07-31 金 5 ファイナルトライアル 50 ピーナッツ 外部記憶ペーパー使用可
e ラーニング予習問題 x2 2015-07-02 木 23:55 締切
非参照 Quiz x 2 2015-07-03 金 5

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