7　7回目講義の補足

7.1　ポアソン分布の平均の計算

定理 $X \sim Po(\lambda)$ ならば、

$$E(X) = \lambda,\ V(X) = \lambda \tag{1}$$

(証明) 次の等式（テイラー展開）

$$e^a = 1 + a + a^2/2 + a^3/3! + a^4/4! + \cdots = \sum_{i=0}^{\infty} \frac{a^i}{i!} \tag{2}$$

を用いる。中辺の第 1 項は $1 = a^0/0!$ と考えると右辺とつながる。

$$\begin{aligned}
E(X) &= \sum_{x=0}^{\infty} x\ e^{-\lambda}\frac{\lambda^x}{x!} = e^{-\lambda}\sum_{x=1}^{\infty} x\ \frac{\lambda^x}{x!} = e^{-\lambda}\sum_{x=1}^{\infty} x\ \frac{\lambda^x}{x!} \qquad (x=0\text{ の項は明らかに }0) && (3)\\
&= e^{-\lambda}\sum_{x=1}^{\infty}\frac{\lambda^x}{(x-1)!} \quad \left(\frac{x}{x!} = \frac{1}{(x-1)!}\right) && (4)\\
&= e^{-\lambda}\sum_{x=1}^{\infty}\frac{\lambda\times\lambda^{x-1}}{(x-1)!} = \lambda e^{-\lambda}\sum_{x=1}^{\infty}\frac{\lambda^{x-1}}{(x-1)!} = \lambda e^{-\lambda}\sum_{i=0}^{\infty}\frac{\lambda^i}{i!} \quad (i = x-1\text{ と置いた}) && (5)\\
&= \lambda e^{-\lambda}\times e^{\lambda} = \lambda && (6)
\end{aligned}$$

最終行でテイラー展開を使った。同様にして、$E\{X(X-1)\} = \lambda^2$ が得られる。これを使うと、

$$V(X) = E\{X(X-1)\} + E(X) - \{E(X)\}^2 \tag{7}$$

から分散が求まる。（証明終）

7.2　連続型確率変数

確率変数 X が連続の値を取りうる場合、X を連続型確率変数と言う。例えば、身長や体重、気温、為替レートなどは連続型確率変数である。この場合、X の値域は

$$\{x \mid A < x < B\},\ \{x \mid A \le x \le B\},\ \{x \mid 0 < x\},\ \{x \mid x\text{ は実数 }\} \tag{8}$$

などという形の区間となり、X の挙動は $P(a \le X \le b)$ の値によって定まる。より正確に定義すれば次の通りである。負の値を取らない関数 $f(x)$ が存在して、どのような実数 a, b に対しても確率 $P(a \le X \le b)$ が図中の斜線部の面積で与えられるとき、すなわち次式

$$P(a \le X \le b) = \int_a^b f(x)\mathrm{d}x\ (-\infty < a \le b < \infty)$$

が成立するとき、X を**連続型確率変数** (continuous random variable)、X の確率分布を**連続型確率分布** (continuous distribution)、関数 $f(x)$ を X の**確率密度関数** (probability density function)、あるいは単に密度関数と言う。

形式上、$f(x)$ の定義域は実数の全体 $(-\infty, \infty)$ とし、X が値を取らない集合の上では $f(x) = 0$ とする。

明らかに $P(-\infty < X < \infty) = 1$ であるから、確率密度関数の全範囲での積分値は 1 である。すなわち、

$$\int_{-\infty}^{\infty} f(x)\mathrm{d}x = 1.$$

また、$P(X \leq a)$ を a の関数とみて、$F(a)$ と書き、X の**分布関数** (distribution function) と言う。

$$F(a) = P(X \leq a) \ (-\infty < a < \infty). \tag{9}$$

もちろん、$F(a) = \int_{-\infty}^{a} f(x)\mathrm{d}x$ であり、これより、F を微分すれば密度関数 f が得られる。

$$F'(a) = f(a). \tag{10}$$

7.3 正規分布

(1) **正規分布の確率計算**: $Z \sim N(0,1)$ とする。

$$P(Z \leq z) = \Phi(z), \quad -\infty < z < \infty \tag{11}$$

と置く。$\Phi(z)$ を標準正規分布 $N(0,1)$ の累積分布関数とする。$\Phi(z)$ の値は、エクセル関数 **normsdist(z)** で計算することが出来る[1]。

基本例：
(*) $P(Z \leq 1.96) = \Phi(1.96) = 0.975, \ P(Z \leq 1.64) = \Phi(1.64) = 0.949 (= 0.95)$
(**) $P(Z \leq 1) = \Phi(1) = 0.841$, $P(Z \leq 2) = \Phi(2) = 0.977, \quad P(Z \leq 3) = \Phi(3) = 0.999$

すぐ分かる通り、

$$P(a < Z) = 1 - P(Z \leq a) = 1 - \Phi(a) \tag{12}$$

である。例えば、$P(1 < Z) = 1 - \Phi(1) = 1 - 0.841 = 0.159.$

$P(Z \leq -1) = \Phi(-1)$ などの値は、標準正規分布 $N(0,1)$ の密度関数の対称性を使って次のように求まる:

$$\Phi(-1) = P(Z \leq -1) = P(Z \geq 1) = 1 - \Phi(1) = 0.159 \tag{13}$$

(2) **公式:** $\Phi(z)$ の値が計算出来るようになれば次の公式を使って計算出来る。

(公式) $Z \sim N(0,1)$ なら $P(a \leq Z \leq b) = \Phi(b) - \Phi(a)$

(発展) $X \sim N(\mu, \sigma^2)$ なら $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$ なので

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

基本例： $Z \sim N(0,1)$ なら、

(1) $P(-1 \leq Z \leq 1) = \Phi(1) - \Phi(-1) = 0.841 - 0.159 = 0.683$

(2) $P(-2 \leq Z \leq 2) = \Phi(2) - \Phi(-2) = 0.977 - 0.023 = 0.954$

(3) $P(-3 \leq Z \leq 3) = \Phi(3) - \Phi(-3) = 0.999 - 0.001 = 0.997$

従って、 $X \sim N(\mu, \sigma^2)$ の場合は次の事実が示せる。$Z = \frac{X-\mu}{\sigma} \sim N(0,1)$ に注意すると、

(1*) $P(\mu - \sigma \leq X \leq \mu + \sigma) = P\left(-1 \leq \frac{X-\mu}{\sigma} \leq 1\right) = P\left(-1 \leq Z \leq 1\right) = (1)$ と同じ

[1] 本テキストのサポートページ（サイエンス社トップページ　http://www.saiensu.co.jp/index.php　で検索欄に「入門統計解析」と入力。テキストのページが開いたら、「サポート情報あり」というボタンをクリック。

(2*) $P(\mu-2\sigma \le X \le \mu+2\sigma) P\left(-2 \le \frac{X-\mu}{\sigma} \le 2\right) = P\left(-2 \le Z \le 2\right) = (2)$ と同じ

(3*) $P(\mu-3\sigma \le X \le \mu+3\sigma) = P\left(-3 \le \frac{X-\mu}{\sigma} \le 3\right) = P\left(-3 \le Z \le 3\right) = (3)$ と同じ

例 7.1. (身長) X を高校生男子の身長とし，$X \sim N(170, 5^2)$ であるならば，X が $[170 \pm 5] = [165, 175]$ の範囲に入る確率はおよそ 68.3%、$[170 \pm 2 \times 5] = [160, 180]$ には 95.4%, $[170 \pm 3 \times 5] = [155, 185]$ には 99.7%であることが分かる.

例 7.2. (在庫量) ある工業部品を出荷する工場を考える。この工場が直面する需要量を X(トン) とおくと、X は正規分布 $N(200, 25^2)$ で表せるとする。品切れとなる確率 0.05 以下となるように在庫量を決めるならば在庫量はどのような範囲となるか。

例 7.3. (コレステロール値) 血液中のコレステロールが多過ぎると心臓病のリスクが高まる。一般に若年の女性の間では、コレステロール過多による病気は少ない。20 歳から 34 歳までの女性のコレステロール値は正規分布に従い、その平均は 185 (mg/dl)、標準偏差は 39 (mg/dl) である。

(1) コレステロール値が 240 (mg/dl) を超えると治療が必要になる。若年女性の何%が治療を必要としているか。

(2) コレステロール値が 200 (mg/dl) を超えると境界域に入ったと見做される。若年女性の何%が、コレステロール値が 200 (mg/dl) から 240 (mg/dl) までの境界域にあるか。

(3) 初老の男性は若年女性に比べてコレステロール値が高くなり易い。55 歳から 64 歳までの男性のコレステロール値は正規分布に従い、平均は 222 (mg/dl)、標準偏差は 37 (mg/dl) である。彼らの内で、高コレステロール（240(mg/dl) 以上）と見做される人は何%いるか。

(4) 初老男性の何%が境界域に入っているか。

7.4 指数分布

定理 1. (指数分布の無記憶性) $X \sim E_X(\lambda)$ とする．このとき，任意の $a, b > 0$ に対して次式が成立する．

$$P(X > a + b | X > b) = P(X > a). \tag{14}$$

逆に (14) が成り立つような $(0, \infty)$ 上の連続型確率分布は指数分布である。

定理 2. (平均と分散) $X \sim E_X(\lambda)$ ならば，X の平均と分散は次のようである．

$$E(X) = 1/\lambda,\ V(X) = 1/\lambda^2,\ D(X) = 1/\lambda. \tag{15}$$

次の例にあるように λ は単位時間当たりの事象の生起回数であり、$1/\lambda$ が事象の生起間隔となる。

例 7.4. (来店間隔) ある会社の電話は 1 時間当たり平均 20 件（$\lambda = 20$(件/時) $= 1/3$(件/分)）である。電話の到着間隔がランダムならば、電話のかかってくる間隔 X(分) は平均 $1/\lambda = 3$(分/件) の指数分布 $E_X(1/3)$ に従う。このとき、15 分以上電話の来ない確率は

$$P(X \ge 15) = e^{-\lambda \times 15} = e^{-5} = 0.0067 \tag{16}$$

と計算できる。6 分以内に来る確率は $P(X \le 6) = 1 - e^{-\lambda \times 6} = 1 - 0.135 = 0.865.\square$

正規分布表

ａに対して、Φ(a)を与える。但し、Φ(a) は標準正規分布の分布関数である。

すなわち、Z～N(0,1) のとき、Φ(a) = P(Z≦a) .　　153頁(4.5.11)式を参照のこと。

a	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.500	0.504	0.508	0.512	0.516	0.520	0.524	0.528	0.532	0.536
0.1	0.540	0.544	0.548	0.552	0.556	0.560	0.564	0.567	0.571	0.575
0.2	0.579	0.583	0.587	0.591	0.595	0.599	0.603	0.606	0.610	0.614
0.3	0.618	0.622	0.626	0.629	0.633	0.637	0.641	0.644	0.648	0.652
0.4	0.655	0.659	0.663	0.666	0.670	0.674	0.677	0.681	0.684	0.688
0.5	0.691	0.695	0.698	0.702	0.705	0.709	0.712	0.716	0.719	0.722
0.6	0.726	0.729	0.732	0.736	0.739	0.742	0.745	0.749	0.752	0.755
0.7	0.758	0.761	0.764	0.767	0.770	0.773	0.776	0.779	0.782	0.785
0.8	0.788	0.791	0.794	0.797	0.800	0.802	0.805	0.808	0.811	0.813
0.9	0.816	0.819	0.821	0.824	0.826	0.829	0.831	0.834	0.836	0.839
1.0	0.841	0.844	0.846	0.848	0.851	0.853	0.855	0.858	0.860	0.862
1.1	0.864	0.867	0.869	0.871	0.873	0.875	0.877	0.879	0.881	0.883
1.2	0.885	0.887	0.889	0.891	0.893	0.894	0.896	0.898	0.900	0.901
1.3	0.903	0.905	0.907	0.908	0.910	0.911	0.913	0.915	0.916	0.918
1.4	0.919	0.921	0.922	0.924	0.925	0.926	0.928	0.929	0.931	0.932
1.5	0.933	0.934	0.936	0.937	0.938	0.939	0.941	0.942	0.943	0.944
1.6	0.945	0.946	0.947	0.948	0.949	0.951	0.952	0.953	0.954	0.954
1.7	0.955	0.956	0.957	0.958	0.959	0.960	0.961	0.962	0.962	0.963
1.8	0.964	0.965	0.966	0.966	0.967	0.968	0.969	0.969	0.970	0.971
1.9	0.971	0.972	0.973	0.973	0.974	0.974	0.975	0.976	0.976	0.977
2.0	0.977	0.978	0.978	0.979	0.979	0.980	0.980	0.981	0.981	0.982
2.1	0.982	0.983	0.983	0.983	0.984	0.984	0.985	0.985	0.985	0.986
2.2	0.986	0.986	0.987	0.987	0.987	0.988	0.988	0.988	0.989	0.989
2.3	0.989	0.990	0.990	0.990	0.990	0.991	0.991	0.991	0.991	0.992
2.4	0.992	0.992	0.992	0.992	0.993	0.993	0.993	0.993	0.993	0.994
2.5	0.994	0.994	0.994	0.994	0.994	0.995	0.995	0.995	0.995	0.995
2.6	0.995	0.995	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.996
2.7	0.997	0.997	0.997	0.997	0.997	0.997	0.997	0.997	0.997	0.997
2.8	0.997	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.998
2.9	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.999	0.999	0.999
3.0	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
3.1	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
3.2	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
3.3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
3.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
3.5	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
3.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
3.7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
3.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
3.9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
4.0	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000