確率・統計（その２）

*統計の基礎（その２）

ベルヌーイ(二項)分布（$B(n,p)$）　$P(x) = {}_nC_x p^x q^{n-x}$，$E[X] = np(=\mu)$，$V[X] = npq$

ポアソン分布（$P_o(\mu)$）　$P(x) = e^{-\mu}\dfrac{\mu^x}{x!}$，$E[X] = \mu$，$V[X] = \mu$

正規分布（$N(\mu,\sigma^2)$）　$f(x) = \dfrac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$，$E[X] = \mu$，$V[X] = \sigma^2$

標準正規分布（$N(0,1)$）　$f(z) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$　（$Z = \dfrac{X-\mu}{\sigma}$により標準化）

χ^2二乗分布

標準正規分布に従う独立な確率変数を$X_1, X_2, \cdots, X_n$とするとき、

$\chi^2 = X_1^2 + X_2^2 + \cdots + X_n^2$は自由度$n$の分布に従う。

確率密度　$k(x,n) = \dfrac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} x^{\frac{\pi}{2}-1} e^{-\frac{1}{2}}$、自由度１のとき　$k(x,1) = \dfrac{1}{\sqrt{2\pi x}} e^{-\frac{1}{2}}$

標本平均　$\overline{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$

中心極限定理

Xが正規分布$N(\mu,\sigma^2)$に従うとき、$\overline{X}$は平均μ、分散$\dfrac{\sigma^2}{n}$の正規分布に従う。

確率変数　$Z = \dfrac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$　は、$n \to \infty$のとき標準正規分布$N(0,1)$に従う。

標本分散　$S^2 = \dfrac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$

普遍分散（普遍標本分散）　$s^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$

t分布（スチューデント分布）　$t = \dfrac{\overline{X}-\mu}{\frac{s}{\sqrt{n}}}$　(自由度$m = n-1$のχ^2分布に従う確率変数X)は、

確率密度$g(t,m) = \dfrac{1}{\sqrt{m\pi}}\dfrac{\Gamma\left(\frac{m+1}{2}\right)}{\Gamma\left(\frac{m}{2}\right)}\left(1+\dfrac{t^2}{m}\right)^{-\frac{m+1}{2}}$に従う。

F分布　$f=\dfrac{X_1/m_1}{X_2/m_2}$　(自由度m_1,m_2のχ^2分布に従う確率変数X_1,X_2)は、

確率密度$h(f,m_1,m_2)=\dfrac{\Gamma\left(\dfrac{m_1+m_2}{2}\right)}{\Gamma\left(\dfrac{m_1}{2}\right)\Gamma\left(\dfrac{m_2}{2}\right)}f^{\frac{m_1}{2}-1}\left(1+\dfrac{m_2}{m_1}f\right)^{-\frac{m_1+m_2}{2}}$に従う。

２変数データ(x_i, y_i)について

$$\sigma_x^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^2=\frac{1}{n}\sum_{i=1}^{n}x_i^2-\overline{x}^2,\ \text{ただし}\ \overline{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$$

$$\sigma_y^2=\frac{1}{n}\sum_{i=1}^{n}(y_i-\overline{y})^2=\frac{1}{n}\sum_{i=1}^{n}y_i^2-\overline{y}^2,\ \text{ただし}\ \overline{y}=\frac{1}{n}\sum_{i=1}^{n}y_i$$

共分散　$\sigma_{xy}=\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})=\dfrac{1}{n}\sum_{i=1}^{n}x_iy_i-\overline{x}\,\overline{y}$

相関係数　$\rho_{xy}=\dfrac{\sigma_{xy}}{\sigma_x\sigma_y}$

回帰直線　$y=ax+b$　$\left(a=\dfrac{\sigma_{xy}}{\sigma_x^2}x,\ b=\overline{y}-a\overline{x}\right)$