統計検定（確率分布）

- **二項分布**
  ベルヌーイ試行をn回行い成功する回数Xが従う確率分布

$$P(X=x) = {}_n\mathrm{C}_x p^x (1-p)^{n-x}$$

- **ポアソン分布**
  nが十分大きく確率pが非常に小さい場合、「np=一定 $(np=\lambda)$」と考える
  ある期間に平均λ回起こる現象が、ある期間にX回おきる確率の分布

$$f(x) = \frac{e^{-\lambda}\lambda^x}{x!}$$

確率変数Xがポアソン分布に従うとき、

$$X\text{の期待値}：E(X) = \lambda\text{、分散}：V(X) = \lambda$$

- **幾何分布**
  成功確率がpである独立なベルヌーイ試行を繰り返すとき、初めて成功するまでの試行回数Xが従う確率分布

$$\text{期待値}：E(x) = \frac{1}{p}, \text{分散}：V(x) = \frac{1-p}{p^2}, \text{確率}：P(X=k) = (1-p)^{k-1}p$$

- **正規分布**
  検定、推定、モデルの作成などに活用される連続型確率分布

$$\text{確率密度関数}：f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (-\infty < x < \infty) \ \text{期待値}：E(X) = \mu, \text{分散}：V(X) = \sigma^2$$

- **標準正規分布**
  正規分布の中で「平均$\mu = 0$、分散$\sigma^2 = 1$」である正規分布

$$\text{標準正規分布}：f(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right) \quad (-\infty < x < \infty)$$

- **指数分布**
  次に何かが起こるまでの期間が従う分布
  ある期間に平均してλ回起こる現象が、次に起こるまでの期間Xが指数分布に従うときの確率密度関数

$$\text{確率密度関数}：f(x) = \{ \lambda e^{-\lambda x} \quad x \geq 0\ 0 \quad x < 0 \ \text{期待値}：E(X) = \frac{1}{\lambda}, \text{分散}：V(X) = \frac{1}{\lambda^2}, \text{累積密度関数}：F(x) = 1 - e^{-\lambda x}$$

- **離散一様分布**
  確率変数Xが離散型である場合に、すべての事象に起こる確率が等しい分布
  確率変数Xが離散一様分布に従うとき、$X=k$となる確率$P(X=k)$は、Nを確率変数Xの取りうる個数とすると

$$\text{確率}：P(X=k) = \frac{1}{N}, \text{期待値}：E(X) = \frac{N+1}{2}, \text{分散}：V(X) = \frac{N^2-1}{12}$$

- **連続一様分布**
  確率変数Xがどのような値でも、その時の確率密度関数$f(x)$が一定の値をとる分布

$$\text{確率密度関数}：f(x) = \{ \tfrac{1}{b-a} \quad (a \leq x \leq b)\ 0 \quad (X < a, b < X) \ \text{期待値}：E(X) = \frac{a+b}{2}, \text{分散}：V(X) = \frac{(b-a)^2}{12} \ \text{累積分布関数}：F(x) = \{ 0 \quad (x < a)\ \tfrac{x-}{b-}$$

標準化データ

ある確立変数Xが平均μ、分散σ^2の正規分布z zは「平均0、分散が1の標準正規分布」

$$\text{標準化}：z \longmapsto \frac{X-\mu}{\sigma}$$

2変数の期待値と分散

2つの確率変数XとYの和$X+Y$、差$X-Y$の期待値は

$$E(X+Y) = E(X) + E(Y) \ E(X-Y) = E(X) - E(Y) \ E(XY) = E(X)E(Y)$$

2つの確率変数XとYの和$X+Y$、差$X-Y$の分散は

$$V(X+Y) = V(X) + V(Y) + 2C_ov(X,Y) \ V(X-Y) = V(X) - V(Y) - 2C_ov(X,Y) \ \text{相関係数}：\frac{\text{共分散}}{\text{標準偏差}}$$