2.1 二値変数

5月23日

鈴木直也

ベルヌーイ分布

二値確率変数 $x \in \{0,1\}$ が1つの場合を考える。
$x=1$ で「表」$x=0$ で「裏」をあらわすとする。
さらにここでこのコインが歪んでいて、表と裏の出る確率が同じとは限らないとしよう。
すると表と裏になる確率はパラメータ μ を用いて

$$p(x=1 \mid \mu) = \mu$$

$$p(x=0 \mid \mu) = 1-\mu$$

ただし $0 \le \mu \le 1$ である。
したがって確率分布は $\mu^x (1-\mu)^{1-x}$ となる。
これはベルヌーイ分布として知られている。

ベルヌーイ分布の平均と分散

またこの分布は

$$\sum_{x=0}^{1} p(x|\mu) = \mu^{0}(1-\mu)^{1-0} + \mu^{1}(1-\mu)^{1-1} = 1$$

よって正規化されていることがわかる。

また、この分布の平均と分散は

$$E[x] = \sum_{x=0}^{1} p(x|\mu)x = \mu$$

$$\mathrm{var}[x] = E[x^2] - E[x]^2 = \mu - \mu^2 = \mu(1-\mu)$$

となる。

尤度関数

ここで x の観測値のデータ集合$D=\{x_1,\ldots,x_N\}$があるとしよう。
$p(x|\mu)$ よりこれらの観測値が得られたと仮定したら、μ の尤度関数が得られる。

$$p(D|\mu)=\prod_{n=1}^{N}p(x_n|\mu)=\prod_{n=1}^{N}\mu^{x_n}(1-\mu)^{1-x_n}$$

頻度主義の考えに従うなら、尤度関数を最大にするようにして μ を推定できる。尤度関数の最大化は対数尤度の最大化と等価なのでベルヌーイ分布の対数尤度関数を考える。

$$\ln p(D|\mu)=\sum_{n=1}^{N}\ln p(x_n|\mu)=\sum_{n=1}^{N}\{x_n\ln\mu+(1-x_n)\ln(1-\mu)\}$$

十分統計量

ここで対数尤度関数がN個の観測値 x_n に対して、観測値の

総和のみの関数型になっている。 *

この和はこの分布の下での、このデータに対する十分統計量

となっている。

統計量の値が決められた条件下で、データが出現する条件付き確率分布が
もはやθ には依らない場合に、この統計量は十分統計量という。

二項分布

次に大きさがNのあるデータ集合のうち、

$x=1$ となる観測値の数mの分布を求めてみる。

これを二項分布という

正規化係数を計算するためN回のコイン投げで表がm回

出るような場合について考えると、次のようにあらわされる。

$$\binom{N}{m}\mu^{m}(1-\mu)^{N-m} \quad \text{ただし} \quad \binom{N}{m}=\frac{N!}{(N-m)!\,m!}$$

二項分布の平均と分散

二項分布の平均と分散は以下のとおりである。

$$E[m] = \sum_{m=0}^{N} m \binom{N}{m} \mu^m (1-\mu)^{N-m} = N\mu$$

$$\begin{aligned}\mathrm{var}[m] &= E\left[m - E[m]^2 \right] \\ &= \sum_{m=0}^{N} (m - E[m])^2 \binom{N}{m} \mu^m (1-\mu)^{N-m} = N\mu(1-\mu)\end{aligned}$$