対数分布

hsjoihs

1. $f(k) = \dfrac{-1}{\ln(1-p)}\dfrac{p^k}{k}$ で p の最尤推定

$$L(p) = \prod_{i=1}^{n} \frac{-1}{\ln(1-p)}\frac{p^{x_i}}{x_i}$$

$$\ln L(p) = \sum_{i=1}^{n}\left(\ln\left(\frac{-1}{\ln(1-p)}\right) + x_i \ln p - \ln x_i\right)$$

$$\frac{\partial \ln L(p)}{\partial p} = \sum_{i=1}^{n}\frac{1}{(1-p)\ln(1-p)} + \sum_{i=1}^{n}\frac{x_i}{p}$$

これを 0 とおくと

$$-\frac{n}{(1-\hat{p})\ln(1-\hat{p})} = \sum_{i=1}^{n}\frac{x_i}{\hat{p}}$$

$$-\frac{n\hat{p}}{(1-\hat{p})\ln(1-\hat{p})} = \sum_{i=1}^{n} x_i$$

$$-\frac{n\hat{p}}{(1-\hat{p})\ln(1-\hat{p})} = n\bar{x}$$

$q = \dfrac{1}{1-\hat{p}}$ とおくと $\hat{p} = 1 - 1/q$ であり、また $q \geq 1$ である。

$$\frac{pq}{\ln q} = \bar{x}$$

$$\frac{q-1}{\ln q} = \bar{x}$$

$r = \ln q$ とおくと $r > 0$ で、$\dfrac{e^r - 1}{r} = \bar{x}$ 。$\bar{x} > 1$ である限りこの解は一意に存在する。

2. W 関数で解く

さて、W 関数で解くことを考えてみよう。$u = \dfrac{1}{\bar{x}}$ とおく。$-s = r + u$ とおけば $e^{-s-u} = -\bar{x}$ なので $-ue^{-u} = se^s$ 。これの $s = -u$ でないほうの解を求めればいい。$-u > -1$ なので $s = W_{-1}\left(-ue^{-u}\right)$ である。具体例として、$\bar{x} = \dfrac{215}{168}$ であるときは、$s = -1.25601$ であり、$r = -s - u = 0.47461$ で $q = e^r = 1.60739$ なので $\hat{p} = 0.37788$ となる。

ここまでの操作は結局、$-\dfrac{\hat{p}}{\left(1-\hat{p}\right)\ln\left(1-\hat{p}\right)} = \bar{x}$ を解いて、$\hat{p}$ を $\bar{x}$ の閉じた式で表す方法を求めるという操作である。後で使うのでまとめておくと、$\hat{p} = 1 - \exp\left(u + W_{-1}\left(-ue^{-u}\right)\right), u = \dfrac{1}{\bar{x}}$ である。ということで、$Q(u) = 1 - \exp\left(u + W_{-1}\left(-ue^{-u}\right)\right)$ を定義すれば $\hat{p} = Q\left(\dfrac{1}{\bar{x}}\right)$ と書ける。

3. estimator の良さ

さて、実は $E(x) = \dfrac{-1}{\ln(1-p)}\dfrac{p}{1-p}$ である。つまりこれは $p = Q\left(\dfrac{1}{E(x)}\right)$ ということであり、ここから直ちに「n が十分大きい時 p の estimator $\hat{p}$ は p の真の値に近づく」ということが分かる。

$E\left(\hat{p}\right)$ とか $Var\left(\hat{p}\right)$ とか求めるのは無理そうで、普通に $\bar{x}$ を攻めればよさそうという話になる。

$E(\bar{x}) = E(x) = \dfrac{-1}{\ln(1-p)}\dfrac{p}{1-p}$、$Var(\bar{x}) = -np\dfrac{p + \ln(1-p)}{(1-p)^2\left(\ln(1-p)\right)^2}$ であるので、中心極限定理より、n が十分大きいときは

$$P\left(\bar{x} - \frac{z_{\alpha/2}}{\sqrt{n}\sqrt{Var(x)}} \leq E(x) \leq \bar{x} + \frac{z_{\alpha/2}}{\sqrt{n}\sqrt{Var(x)}}\right) = 1 - \alpha$$

である。ここで $Var(x)=-p\dfrac{p+\ln(1-p)}{(1-p)^2\bigl(\ln(1-p)\bigr)^2}$ の p を $\hat{p}$ で推定することを考えると、まずは $-\dfrac{\hat{p}}{(1-\hat{p})\ln(1-\hat{p})}=\bar{x}$ なので $-\bar{x}^2+\dfrac{\bar{x}}{1-\hat{p}}$ と書け、つまり $-\bar{x}^2+\dfrac{\bar{x}}{\exp\bigl(u+W_{-1}(-ue^{-u})\bigr)}$