Confidence interval on the cross entropy

Jean-Michel Marin

February 2022

CNRS – IMAG (Montpellier, France)

Goal

- True distribution F (density f)
- We observed an iid sample $(y_1, \ldots, y_n)$ from F
- We considered the parametric model Q_θ (density $q(y|\theta)$)
- The cross-entropy is

$$\beta = \max_\theta \mathbb{E}\left[\log(q(y|\theta)\right] = \max_\theta \log(q(y|\theta))f(y)\mathrm{d}y$$

Construct a confidence interval on β

Estimate of β

- Clearly, $\beta = \gamma(F)$
- $\hat{F}_n$ the empirical distribution associated to $(y_1, \dots, y_n)$
- $\hat{\theta}_n$ the MLE of θ
- An estimate of β is

$$\hat{\beta}_n = \gamma(\hat{F}_n) = \frac{1}{n}\sum_{i=1}^{n} \log(q(y_i|\hat{\theta}_n))$$

We propose to bootstrap the statistics $S = \hat{\beta}_n - \beta$

A gaussian toy example $F = \mathcal{N}(0,1)$; $Q_\theta = \mathcal{N}(\theta,1)$

In such a case

- $\beta = -1/2 \log(2\pi) - 1/2$
- $\hat{\theta}_n = \bar{y}_n$; $\hat{\beta}_n = -1/2 \log(2\pi) - 1/(2n) \sum_{i=1}^n (y_i - \bar{y}_n)^2$
- The bootstrap version of $S = \hat{\beta}_n - \beta = \gamma(\hat{F}_n) - \gamma(F)$ is

$$S^* = 1/(2n) \sum_{i=1}^n (y_i - \bar{y}_n)^2 - 1/(2n) \sum_{i=1}^n (y_i^* - \bar{y}_n^*)^2 = 1/2 s_n^2 - 1/2 s_n^*$$

where $(y_1^*, \dots, y_n^*)$ is a bootstrap sample from $(y_1, \dots, y_n)$

Numerical experiments for the gaussian toy example $F = \mathcal{N}(0,1)$; $Q_\theta = \mathcal{N}(\theta,1)$

```r
beta <- -1/2*log(2*pi)-1/2
N <- 1000 ; n <- 500 ; B <- 500 ; alpha <- 0.1
statboot <- rep(0,B) ; cover <- 0
for (i in 1:N) {
  x <- rnorm(n) ; sn2 <- var(x)*(n-1)/n
  for (j in 1:B) {
    xstar <- sample(x,n,rep=TRUE) ; sn2star <- var(xstar)*(n-1)/n
    statboot[j] <- sn2/2-sn2star/2 }
  interva <- -1/2*log(2*pi)-sn2/2-quantile(statboot,c(1-alpha/2,alpha/2),
  names=FALSE)
  cover <- cover + (beta >= interva[1] & beta <= interva[2]) }
cover/N
```

we get 0.903

A gaussian example with unknown variance ; $F = \mathcal{N}(0,1)$; $Q_{(\mu,\sigma^2)} = \mathcal{N}(\mu,\sigma^2)$

In such a case

- $\beta = -1/2 \log(2\pi) - 1/2$
- $\hat{\theta}_n = \left(\bar{y}_n, s_n^2\right)$; $\hat{\beta}_n = -1/2(\log(2\pi) + 1) - 1/2 \log(s_n^2)$
- The bootstrap version of $S = \hat{\beta}_n - \beta = \gamma(\hat{F}_n) - \gamma(F)$ is

$$S^* = 1/2 \log(s_n) - 1/2 \log(s_n^*)$$

Numerical experiments for the gaussian toy example $F = \mathcal{N}(0,1)$; $Q_{(\mu,\sigma^2)} = \mathcal{N}(\mu, \sigma^2)$

```
beta <- -1/2*log(2*pi)-1/2
N <- 1000 ; n <- 500 ; B <- 500 ; alpha <- 0.1
statboot <- rep(0,B) ; cover <- 0
for (i in 1:N) {
  x <- rnorm(n) ; sn2 <- var(x)*(n-1)/n
  for (j in 1:B) {
    xstar <- sample(x,n,rep=TRUE) ; sn2star <- var(xstar)*(n-1)/n
    statboot[j] <- log(sn2)/2-log(sn2star)/2 }
  interva <- -1/2*log(2*pi)-sn2/2-quantile(statboot,c(1-alpha/2,alpha/2),
  names=FALSE)
  cover <- cover + (beta >= interva[1] & beta <= interva[2]) }
cover/N
```

we get 0.896