베이지안 절차

정의

- 개요

1. 베이지안 정리에서 $\frac{g(\theta)g(y|\theta)}{\sum g(\theta)g(y_i|\theta)}$는

1) 사전확률 θ가 주어졌을 때

2) y라는 추가 정보가 주어졌을 경우 변동되는 사후 확률로 정의할 수 있으며

3) 이를 모든 확률변수로 확대하면 다음과 같은 꼴이 도출된다.

$$(1)\frac{g(\theta_i)g(y|\theta_i)}{\sum_{j=1}^{n} g(\theta_j)g\left(y|\theta_j\right)}(i=1,2,\cdots,n) \Rightarrow \frac{h(\theta)L(y|\theta)}{\int_{-\infty}^{\infty} h(\theta)g(y|\theta)\,d\theta} = k(\theta|y)$$

2. 즉, 베이지안 확률론을 모수의 추정에 대입한 것

1) 모수에 대한 확률변수 Θ에 대한 pdf $h(\theta)$ 와

2) 그 분포에 모수를 의존하는 확률변수 X의 pdf $L(X|\theta)$ 결합하여 $h(\theta)L(X|\theta)$

3) 모수에 대해 적분한 후 $\int_{-\infty}^{\infty} h(\theta)g(y|\theta)\,d\theta$

4) 그것을 결합 pdf $h(\theta)L(X|\theta)$ 에 대해 $\int_{-\infty}^{\infty} h(\theta)g(y|\theta)\,d\theta$ 로 통제하면 베이지안 사후분포가 나온다.

정의

- 사전분포와 사후분포

1. 위에서, 모수에 대한 $h(\theta)$ 를 pdf로 가지는 확률변수 θ의 분포를 사전분포라고 한다.
2. 위에서, $\frac{h(\theta)L(y|\theta)}{\int_{-\infty}^{\infty} h(\theta)g(y|\theta)\,d\theta} = k(\theta|y)$로 도출되는 $k(\theta|y)$를 pdf로 가지는 분포를 사후분포라고 한다.
3. 한편, 사후 분포는 θ에 의존하지 않는 '상수' C(y)와 θ에 의존하는 항의 인수로 분해될 수 있다.

1) 즉, $k(\theta|y) = \frac{h(\theta)L(y|\theta)}{\int_{-\infty}^{\infty} h(\theta)g(y|\theta)\,d\theta} = \mathrm{c}(\mathrm{y})h(\theta)L(y|\theta)$ 이고, 만약 상수를 고려하지 않는다면

2) $k(\theta|\mathrm{y}) \propto h(\theta)L(y|\theta)$ 의 <u>비례 관계</u>로 정의할 수 있다.

3) 한편, 만약 $Y = U(X)$인 X에 대한 충분통계량이라면, 이는 또한 $h(\theta)L(u(x)|\theta)$ 로 나타낼 수 있다.

정의

- 켤레분포족(공액분포)

1. 사전 분포의 pdf와 표본 분포의 pdf를 결합하였을 때 도출되는 사후 분포가 사전 분포와 동일한 분포족에 속한다면
2. 이를 표본 분포 X에 대한 사전 분포가 켤레분포족에 속한다고 정의한다.

1) 다시 말해, $k(\theta|y) \propto L(x|\theta)h(\theta)$를 정의할 때
2) 만약 $k(\theta|y)$의 분포가 $h(\theta)$의 분포와 동일한 분포족에 속할 때
3) $h(\theta)$의 분포를 $L(x|\theta)$가 따르는 분포에 대한 켤레분포족이라고 정의한다.

표본의 분포	사전 분포	사후분포
$Bin(n,\theta)$	$\theta \sim beta(a,\beta)$	$beta(c,d)$
$poi(\theta)$	$\theta \sim gamma(a,\beta)$	$gamma(c,d)$
$N(\theta,\sigma^2)$ (단, 모분산은 알려져있다.)	$\theta \sim N(a,\beta)$	$N(c,d)$
$N(\mu,\theta)$ (단, 모평균은 알려져있다.)	$\theta \sim \frac{1}{gamma(a,\beta)}$	$gamma(a,\beta)$

정의

- 베이지안 점추정

1. 위험함수 $E[\mathcal{L}\{\Theta,\delta(X)\}|X = x]= \int_{-\infty}^{\infty}\mathcal{L}\{\Theta,\delta(X)\}k(\theta|x)\,d\theta$를 정의하자.

1) 이 때, 이 위험함수값을 최소화하는 추정량 $\delta(X)$를 베이지안 추정량이라고 한다.

2) 만약 손실함수가 절대손실오차함수(MAE)로 주어지는 경우

(1)$\delta(X)$ = $\mathrm{E}[\Theta|X]$ 일 때 위험함수가 최소화되므로, 베이지안 추정값은 중위수가 된다.

3) 만약 손실함수가 제곱손실오차함수(MSE)로 주어지는 경우

(1) $\delta(X)$ = $\mathrm{E}[\Theta|X]$ 일 때 위험함수가 최소화되므로, 베이지안 추정값은 평균값이 된다.

정의

- 베이지안 점추정

2. 위험함수의 기댓값은 다음과 같이 구할 수 있다.

1) $\mathrm{g(x)} = \int_{-\infty}^{\infty} h(\theta)g(x|\theta)\,d\theta$ 라고 할 때

(1) $\mathrm{E}\left[E[\mathcal{L}\{\Theta,\delta(X)\}|X=x]\right] = \int_{-\infty}^{\infty}\{\int_{-\infty}^{\infty}\mathcal{L}\{\Theta,\delta(X)\}k(\theta|x)\,d\theta\}\,\mathrm{g(x)}\,dx$

$$= \int_{-\infty}^{\infty}\{\int_{-\infty}^{\infty}\mathcal{L}\{\Theta,\delta(X)\}\frac{h(\theta)L(x|\theta)}{\cancel{\int_{-\infty}^{\infty}h(\theta)g(x|\theta)\,d\theta}}\,d\theta\}\cancel{\int_{-\infty}^{\infty}h(\theta)g(x|\theta)\,d\theta}\,dx$$

$$= \int_{-\infty}^{\infty}\{\int_{-\infty}^{\infty}\mathcal{L}\{\Theta,\delta(X)\}L(x|\theta)\,dx\}\,h(\theta)\,d\theta$$

(2) 이 때, $\delta(X)$를 극소화하면 위험의 평균을 최소화하게 된다.

3. 한편, 이 베이지안 추정해가 $\delta(X)$에 대한 항과 나머지 항의 선형결합으로 표현되어서

1) $n \to \infty$일 때 어떤 특정값으로 '수렴(축소)'하는 관계를 보여준다면

2) 이를 축소 추정값 이라고 표현한다.

정의

- 베이지안 구간 추정

1. θ의 구간추정을 원할 경우

1) $\alpha = \mathrm{P}[u(x) < \Theta < v(x) | X = x] = \int_{u(x)}^{v(x)} k(\theta|x)\, d\theta$ 인 $v(x)$와 $u(x)$를 찾는다.

2) 이를 <u>신용구간</u> 혹은 <u>확률구간</u>이라고 표현하며, 신뢰구간과는 엄연히 다른 개념이다.

정의

- 베이지안 검정절차

1. H_0 : $\theta \epsilon w_0$ VS H_1 : $\theta \epsilon w_1$의 가설을 검정한다고 할 때

1) 각각의 가설에 대한 확률 즉

$(1) P(\Theta \epsilon w_0|X) = P_{H_0}$

$(2) P(\Theta \epsilon w_1|X) = P_{H_1}$

2) 이 때

(1) $P_{H_1} > P_{H_0}$ 이면 H_1을 채택한다.

(2) $P_{H_0} > P_{H_1}$ 이면 H_0를 채택한다.

2. 특이한 점은, 두 개 이상의 가설에 대해 동시적으로 검정을 실시할 수 있다는 점이다.

1) 왜냐하면 단순히 조건부 확률이 더 큰 쪽을 옳은 가설로 선택하기 때문이다.

2) 따라서, 동시 검정하고자 하는 가설을 조건부 확률에 포함하여 검정한다.

예제

- 베이지안 절차

1. $X_i|\theta \sim \mathrm{N}(\theta, \sigma^2)$ 이고 $\Theta \sim \mathrm{N}(\theta_0, {\sigma_0}^2)$ 이라 하자.

2. $Y = \bar{x}$ 인 충분통계량일 때, 결론적으로 모수 θ에 대한 베이지안 점추정은

1) $k(\theta|y) \propto \exp\left[\frac{\left({\sigma_0}^2 + {\sigma^2}/_{\mathrm{n}}\right)\theta^2 - 2\left(\mathrm{y}{\sigma_0}^2 + \theta_0\left({\sigma^2}/_{\mathrm{n}}\right)\right)\theta}{2\left(\frac{\sigma^2}{n}\right){\sigma_0}^2}\right]$ 의 비례관계가 성립된다.

(1) 위를 정리하면 $k(\theta|y) \propto \exp\left[\frac{\left(\theta - \frac{\mathrm{y}{\sigma_0}^2 + \theta_0\left({\sigma^2}/_{\mathrm{n}}\right)}{{\sigma_0}^2 + \frac{{\sigma_0}^2}{n}}\right)^2}{2\frac{\left({\sigma^2}/_{\mathrm{n}}\right){\sigma_0}^2}{{\sigma_0}^2 + {\sigma^2}/_{\mathrm{n}}}}\right]$ 이고

(2) $\mathrm{N}\left[\frac{\mathrm{y}{\sigma_0}^2 + \theta_0\left({\sigma^2}/_{\mathrm{n}}\right)}{{\sigma_0}^2 + \frac{{\sigma_0}^2}{n}}, \frac{\left({\sigma^2}/_{\mathrm{n}}\right){\sigma_0}^2}{{\sigma_0}^2 + {\sigma^2}/_{\mathrm{n}}}\right]$인 정규분포에서, 손실함수가 MSE로 주어질 때

- 위험함수를 최소화하는 베이지안 추정해는 이 분포의 평균이 된다.
- 따라서 $\theta = \delta(Y) = \frac{\mathrm{y}{\sigma_0}^2 + \theta_0\left({\sigma^2}/_{\mathrm{n}}\right)}{{\sigma_0}^2 + \frac{\sigma^2}{n}} = \left(\frac{{\sigma_0}^2}{{\sigma_0}^2 + \frac{\sigma^2}{n}}\right)y + \frac{\theta_0\left({\sigma^2}/_{\mathrm{n}}\right)}{{\sigma_0}^2 + \frac{\sigma^2}{n}}$ 은 베이지안 추정해가 된다.
- 이 때, 가장 오른쪽의 표현형은 축소추정값을 표현하는 형태이며, 는 $n \to \infty$ 일 때 점차 y로 수렴한다.

예제

- 베이지안 검정절차

1. $H_0 : \theta \leq 10$ VS $H_1 : \theta > 10$의 가설을 검정한다고 하자.

2. 사전분포로 $\Gamma(a = 10, \beta = 1.2)$ 를 가정했다고 하고,

   이 때 표본은 푸아송 분포에서 무작위로 추출되었다고 하자.

1) 이 때, 푸아송분포의 켤레분포는 감마분포로, 이 두 분포의 결합은

   같은 감마분포를 도출하게 된다.

2) 이 때

(1) $k(\theta|y) \propto \mathrm{pdf}\left\{f : f(x)\text{는 } \Gamma\left(\sum x_i + a, \frac{\beta}{n\beta+1}\right)\text{의 } pdf\right\}$ 와 비례관계에 있고

(2) $\sum x_i$ = 177, n = 20이라고 한다면 사후분포는 $\Gamma(187{,}0.48)$로 변화한다.

3) 사후 분포로 가설을 검정하면

(1) $p_{H_0} = \mathrm{P}[\Gamma(187{,}0.48) \leq 10]$ = 0.9368

(2) $p_{H_1} = \mathrm{P}[\Gamma(187{,}0.48) > 10]$ = 1- 0.9368

4) 따라서 H_0를 채택한다.