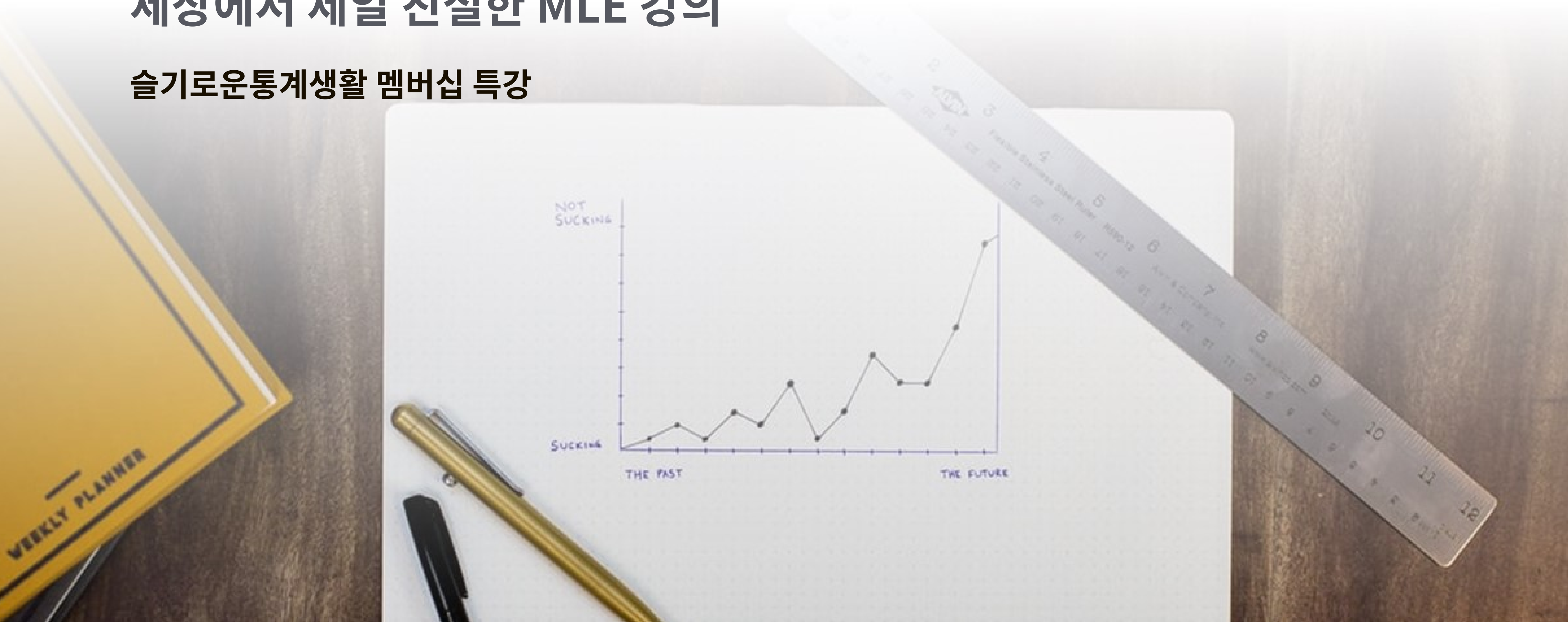
나는야 통계왕이 될꺼야!

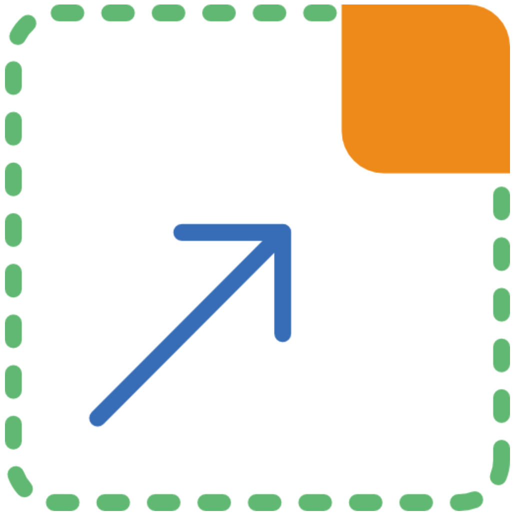
세상에서 제일 친절한 MLE 강의

슬기로운통계생활 멤버십 특강



최대우도 추정량

Maximum Likelihood Estimator



모수를 합리적으로 추정하는 방법

확률변수의 모수

확률변수는 모수에 의하여 그 성질이 결정된다. 문제는 **모수를 찾는 방법**이 관건이다.

- 베르누이 확률변수: 성공확률 - p
- 균일분포 확률변수: 최소값 - a, 최대값 - b
- 포아송분포 확률변수: 평균값 - λ

가장 그럴듯한 모수는 무엇일까?

균일분포를 따르는 확률변수 X의 최소값은 0, 최대값은 θ로 알려져있다.

$$X \sim U(0, \theta)$$

이러한 확률변수 X에서 표본을 5 개를 뽑았을때, 다음의 관찰값이 나 왔다고 하자.

6.469, 3.942, 6.185, 4.769, 1.361

다음 중 가장 그럴듯한 θ 값은?

1) -1

2) 3

3) 8

4) 1000

최대우도 추정량

표본들로 미루어보아, 우리가 가진 표본들을 (고정) 발생시킬 수 있는 **가능성이 가장 높은** 값을 모수를 **추정**한다.

$$6.469, 3.942, 6.185, 4.768, 1.361$$

1) -1 : 균일분포 정의에 따라 성립불가.

2) 3: 표본이 3보다 크므로 가능성이 없음.

3) 8: 가능성이 있음.

4) 1000: 가능성이 있으나, 아주 희박함.

가장 가능성이 있는 것은 8. 왜 그럴까?

균일분포 $U(0, \theta)$ 의 가능도 함수

$$
\begin{aligned}
L\left(\theta; \underline{x}\right) &= \Pi_{i=1}^{n} \left(\frac{1}{\theta}\right) \mathbb{1}_{(0 \le x_i \le \theta)} \\
&= \left(\frac{1}{\theta}\right)^n \mathbb{1}_{(0 \le min(\underline{x}))} \mathbb{1}_{(max(\underline{x}) \le \theta)} \\
&\propto \left(\frac{1}{\theta}\right)^n \mathbb{1}_{(max(\underline{x}) \le \theta)}
\end{aligned}
$$

- θ가 8일 경우 각 표본의 가능도 (Likelihood)

$$L(8; \underline{x}) = \left(\frac{1}{8}\right)^4$$

- θ가 3일 경우 각 표본의 가능도 (Likelihood)

$$L(3; \underline{x}) = 0$$

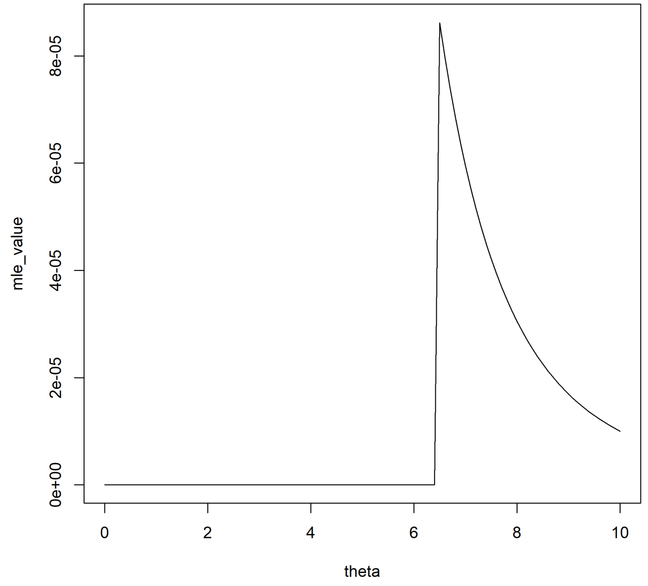
언제가 최대가 될까?

가능도 함수가 최대가 되는 지점은 표본의 최대값에서 최대가 됨.

즉, **최대우도추정'치'** 는 관찰값의 최대값 6.469가 된다.

최대우도 추정량은 모수 θ를 추정하는 방법을 의미함. 즉, 다른 데이터 셋이 주어져도, 그것들의 최대값을 사용해서 θ를 추정하는 방법을 의미한다.

$$\hat{\theta} = max(\underline{X})$$



베르누이 분포의 모수 p

확률변수 X가 베르누이 확률변수 모수 p를 따를때, 다음의 표본이 관찰되었다. 최대우도 추정량을 구해보자.

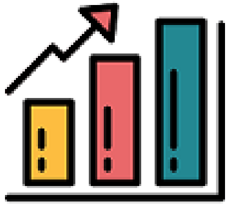
$$0, 1, 1, 0, 0, 1, 0, 0, 1, 1$$

가능도 함수

$$\begin{aligned} L\left(p; \underline{x}\right) &= \Pi_{i=1}^{n} p^{x_i} \left(1-p\right)^{1-x_i} \mathbb{1}_{\left(x_i \in \{0,1\}\right)} \\ &= p^{\sum x_i} \left(1-p\right)^{n-\sum x_i} h\left(\underline{x}\right) \\ &= p^{x_+} \left(1-p\right)^{n-x_+} h\left(\underline{x}\right) \end{aligned}$$

수식에서 x_+는 $\sum x_i$를 의미한다.

로그가능도 함수 (Log Likelihood function)



우리는 가능도 함수를 최대로 만드는 값에 관심이 있으므로, 로그 함수를 취해도 MLE 값은 변하지 않음.

$$\ell\left(p;\underline{x}\right) = x_{+} log\left(p\right) + \left(n - x_{+}\right) log\left(1 - p\right) + c\left(\underline{x}\right)$$

로그 가능도 함수를 최대로 만드는 p 값은 기울기가 0이 되는 조건을 만족해야함.

스코어 함수 (Score function)

로그 가능도 함수를 모수에 대하여 미분한 그레디언트 함수를 0값으로 만드는 p를 구한다.

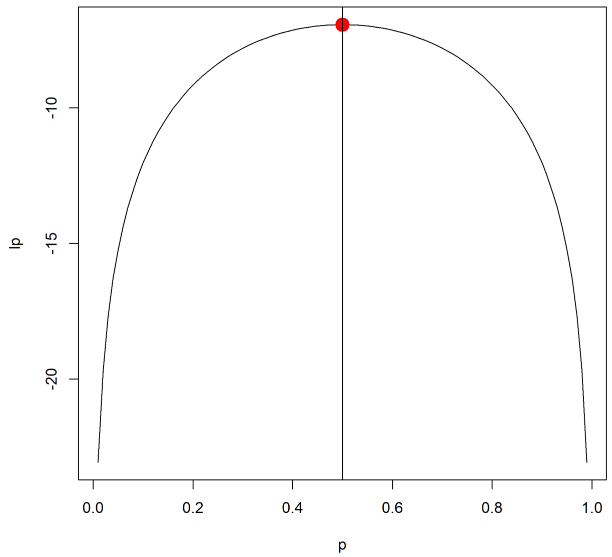
$$\frac{\partial \ell}{\partial p} = \frac{x_+ - np}{p(1-p)} \overset{set}{=} 0$$

위의 식을 풀어보면, 다음과 같다.

$$\hat{p} = \frac{x_+}{n} = \frac{\sum x_i}{n} = \bar{x}$$

즉, 최대우도 추정량은 다음과 같다.

$$\hat{p} = \overline{X}$$

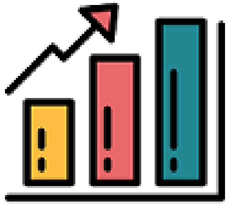


다음시간

MLE 중급

- MLE의 신뢰구간을 구하기위한 여정
- Information matrix
- MLE의 점근 분포

슬기로운 통계생활 강의 및 멤버십 홍보



클래스101 엑셀로 배우는 기초 통계

- 엑셀로 시작하는 기초통계 강의 (유료).
- 프로그래밍에 친숙하시지 않은 분들을 위해서 엑셀을 사용하다 보니 좀 더 눈에 보이고 직관적입니다. (확률변수와 기술통계, 정규분포, 통계 추론과 검정에 대한 원리 등을 다룹니다!)

유료 멤버십 가입 - 데이터 사이언스 관련 유료 영상들을 한 달 만2천원으로 저렴하게!