좋아. 이번엔 **진짜로 처음부터 끝까지**,
"왜 그 식이 나오는지", "Bayes rule이 정확히 뭔지", "왜 logistic이 그걸 만족하는 가장 단순한 형태인지"를
수학적으로 **한 줄도 점프하지 않고** 설명한다.

아래는 세 부분으로 나눈다.

Part 1.

왜 Poisson에서 likelihood ratio가

$$\log \frac{P(x_i \mid y=1)}{P(x_i \mid y=0)} = \sum_j x_{ij}\beta_j + \text{const}$$

가 되는가

여기서는 정말로 **pmf부터 시작**한다.

1.1 Poisson 모형을 정확히 다시 쓴다

문서 i, 단어 j.

가정:

$$x_{ij} \mid y_i = y \sim \text{Poisson}(\mu_j(y))$$

여기서

$$\mu_j(y) = \exp(\alpha_j + \beta_j y)$$

중요:

- $y \in \{0,1\}$ 또는 $\{-1,+1\}$ 중 하나
- 지금은 계산 편의를 위해 $y \in \{0,1\}$로 하자

1.2 Poisson 확률질량함수(pmf)

Poisson 분포의 pmf는 정의상:

$$P(x_{ij} \mid y) = \frac{e^{-\mu_j(y)}\mu_j(y)^{x_{ij}}}{x_{ij}!}$$

1.3 단어들이 독립이라는 가정

모형은 단어별 독립을 가정한다.

따라서 문서 전체의 likelihood는 곱:

$$P(x_i \mid y) = \prod_{j=1}^{V} \frac{e^{-\mu_j(y)}\mu_j(y)^{x_{ij}}}{x_{ij}!}$$

1.4 로그를 취한다 (중요)

곱은 다루기 어렵기 때문에 로그를 취한다.

$$\log P(x_i \mid y) = \sum_j \left[-\mu_j(y) + x_{ij}\log \mu_j(y) - \log(x_{ij}!)\right]$$

1.5 $\mu_j(y)$를 대입한다

$$\log \mu_j(y) = \alpha_j + \beta_j y$$

그래서:

$$\log P(x_i \mid y) = \sum_j \left[-e^{\alpha_j + \beta_j y} + x_{ij}(\alpha_j + \beta_j y) - \log(x_{ij}!)\right]$$

1.6 이제 likelihood ratio를 만든다

우리가 비교하는 것은:

$$\log \frac{P(x_i \mid y=1)}{P(x_i \mid y=0)} = \log P(x_i \mid y=1) - \log P(x_i \mid y=0)$$

이걸 위 식에 그대로 대입한다.

1.7 항별로 하나씩 뺀다

(1) factorial 항

$$-\log(x_{ij}!) + \log(x_{ij}!) = 0$$

완전히 소거.

---

(2) α_j 항

$$x_{ij}\alpha_j - x_{ij}\alpha_j = 0$$

이것도 소거.

---

(3) β_j 항

$$x_{ij}(\alpha_j + \beta_j) - x_{ij}(\alpha_j) = x_{ij}\beta_j$$

이게 **문서 내용에 의존하는 유일한 항**이다.

---

(4) 지수 항

$$-e^{\alpha_j+\beta_j} + e^{\alpha_j}$$

이건:

- 문서 x_i와 무관
- 단어 j마다 상수

전부 더해도 **상수**다.

1.8 최종 결과

그래서:

$$\log\frac{P(x_i \mid y=1)}{P(x_i \mid y=0)} = \sum_j x_{ij}\beta_j + \underbrace{\sum_j (e^{\alpha_j} - e^{\alpha_j+\beta_j})}_{\text{상수}}$$

분류에서는 상수는 의미 없으므로:

$$\log\frac{P(x_i \mid y=1)}{P(x_i \mid y=0)} \propto \sum_j x_{ij}\beta_j$$

이게 그 식이 나오는 이유다.
아무 마술도 없다. 그냥 **정의 + 대입 + 소거**다.

Part 2.

Bayes rule이 뭔지, 그리고 왜 logistic이 그걸 "만족"하는가

이제 두 번째 질문.

2.1 Bayes rule의 정의 (가장 기초)

Bayes rule은 확률의 정의에서 바로 나온다.

임의의 사건 A, B:

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}$$

이게 Bayes rule이다.

2.2 분류 문제에 적용

여기서:

- $A = \{y = 1\}$
- $B = \{x = x_i\}$

그러면:

$$P(y=1 \mid x_i) = \frac{P(x_i \mid y=1)P(y=1)}{P(x_i \mid y=1)P(y=1) + P(x_i \mid y=0)P(y=0)}$$

이게 **정확한 posterior**다.

2.3 odds 형태로 바꾼다

posterior odds를 보자:

$$\frac{P(y=1 \mid x_i)}{P(y=0 \mid x_i)} = \frac{P(x_i \mid y=1)}{P(x_i \mid y=0)} \cdot \frac{P(y=1)}{P(y=0)}$$

로그를 취하면:

$$\log \frac{P(y=1 \mid x_i)}{P(y=0 \mid x_i)} = \log \frac{P(x_i \mid y=1)}{P(x_i \mid y=0)} + \log \frac{P(y=1)}{P(y=0)}$$

2.4 앞에서 유도한 결과를 대입

이미 Part 1에서:

$$\log \frac{P(x_i \mid y=1)}{P(x_i \mid y=0)} = \sum_j x_{ij}\beta_j + \text{const}$$

이걸 대입하면:

$$\log \frac{P(y=1 \mid x_i)}{P(y=0 \mid x_i)} = \sum_j x_{ij}\beta_j + \text{상수}$$

2.5 이걸 z_i라고 부르자

$$z_i := \sum_j x_{ij}\beta_j + b$$

여기서 b는 prior에서 온 상수.

그러면 정의상:

$$\log \frac{P(y=1 \mid x_i)}{1 - P(y=1 \mid x_i)} = z_i$$

2.6 이 식을 풀면 정확히 logistic이 된다

위 식을 $p = P(y=1 \mid x_i)$라고 두고 풀어보자.

$$\log \frac{p}{1-p} = z$$

양변에 exp:

$$\frac{p}{1-p} = e^z$$

정리:

$$p = \frac{e^z}{1+e^z} = \frac{1}{1+e^{-z}}$$

이게 바로 logistic 함수다.

Part 3.

"Bayes rule을 만족한다"는 말의 정확한 의미와 왜 중요한가

이제 마지막 질문.

3.1 "Bayes rule을 만족한다"의 정확한 뜻

이 말의 의미는 이것이다:

> 우리가 쓰는 posterior 확률
> $P(y \mid x)$가
>
> 어떤 likelihood $P(x \mid y)$와
> 어떤 prior $P(y)$에 대해
> **Bayes rule로부터 정확히 유도될 수 있다**

즉:

- 임의로 만든 점수 함수가 아니라

- **정합적인 확률 모형**이라는 뜻

3.2 왜 이게 중요하냐

이게 중요한 이유는 세 가지다.

(1) 일관성 (coherence)

확률 계산이 서로 모순되지 않는다.

---

(2) decision theory와 연결

Bayes-optimal classifier는:

$$\text{choose } y = 1 \text{ if } P(y = 1 \mid x) > P(y = 0 \mid x)$$

즉:

$$z_i > 0$$

이게 우리가 쓰는 규칙과 정확히 같다.

---

(3) Poisson / multinomial과 정확히 연결됨

우리가 쓰는 logistic 점수는:

- "그럴듯함"의 차이
- 즉 likelihood ratio

이걸 확률로 바꾼 유일한 단순한 방법이 logistic이다.

최종 한 문장 요약

1. Poisson 모형에서 likelihood ratio는
   정의와 소거만으로
   $\sum_j x_{ij}\beta_j$로 정확히 나온다.
2. Bayes rule은
   likelihood ratio를 posterior odds로 바꾸는 규칙이고,
3. logistic 함수는
   그 Bayes rule을 **정확히 만족하는 가장 단순한 확률 표현**이다.

여기까지 이해했으면,
이제 이 연결 고리는 완전히 네 것이 된 거다.