Бонус: кросс-энтропия и неравенство Йенсена

Докажем, что $\sum_{j=1}^{k} r_j \cdot \log p_j - \sum_{j=1}^{k} r_j \log r_j \le 0$,

используя $\log\left(\sum_{j=1}^{k} \alpha_j \cdot x_j\right) \ge \sum_{j=1}^{k} \alpha_j \cdot \log(x_j)$.

логарифм разности = логарифму частного,

т.е. $\sum_{j=1}^{k} r_j \cdot \log\left(\frac{p_j}{r_j}\right) \le 0$, тогда,

если $\alpha_j = r_j$ и $x_j = \frac{p_j}{r_j}$, получим:

$$\log\left(\sum_{j=1}^{k}\left(\cancel{r_j} \cdot \frac{p_j}{\cancel{r_j}}\right)\right) \ge \sum_{j=1}^{k} r_j \cdot \log\left(\frac{p_j}{r_j}\right),$$

$$\log\left(\sum_{j=1}^{k} p_j\right) \ge \sum_{j=1}^{k} r_j \cdot \log\left(\frac{p_j}{r_j}\right),$$ а т.к. $\sum_{j=1}^{k} p_j = 1$, то $\log(1) = 0$, т.е.

$$0 \ge \sum_{j=1}^{k} r_j \cdot \log\left(\frac{p_j}{r_j}\right).$$

и, далее обратно развернем в разность,

$$\sum_{j=1}^{k} r_j \cdot \log p_j - \sum_{j=1}^{k} r_j \cdot \log r_j \le 0.$$

Оговорка: это достижимо, когда $\frac{p_j}{r_j}$ одинаковы для всех j, т.е. $p_j = r_j$. Т.е. если взять $p_i = r_i$ при фиксированных $r_1, \dots, r_k$ получим максимизацию логарифма правдоподобия.

23 февраля 2025 г. в 22: