تمرین سری اول تئوری مبانی علم داده

پارسا آقاعلی

۴۰۰۵۲۱۰۷۲

سوال ۱)

حداکثر درست‌نمایی یا **Maximum Likelihood Estimation (MLE)** روشی آماری است که برای تخمین پارامترهای یک توزیع احتمالی بر اساس داده‌های مشاهده‌شده استفاده می‌شود. ایده اصلی MLE این است که پارامترهایی را پیدا کنیم که احتمال وقوع داده‌های مشاهده‌شده را به حداکثر برسانند. در این روش، تابع درست‌نمایی (Likelihood Function) ساخته می‌شود که نشان‌دهنده احتمال داده‌ها تحت فرض پارامترهای مختلف توزیع است. سپس، پارامترهایی را انتخاب می‌کنیم که این تابع درست‌نمایی را به بیشترین مقدار ممکن برساند.

MLE کاربرد گسترده‌ای در یادگیری ماشین، آمار و علوم داده دارد، چرا که بر اساس داده‌های واقعی عمل کرده و پارامترهای مدل را به صورت بهینه تعیین می‌کند. در روشMLE ، از مفهوم **مشتق‌گیری** نیز برای یافتن نقطه ماکسیمم تابع درست‌نمایی استفاده می‌شود. این روش به‌ویژه در مدل‌های خطی و توزیع‌های شناخته‌شده مانند نرمال، پواسون و نمایی کاربرد دارد.

سوال ۲)

$$\log L(\theta) = \sum_{i=1}^{i=n} \log\left(\frac{-\theta x_i^{\theta-1}}{2^{\theta}}\right)$$

$$\log L(\theta) = \sum_{i=1}^{i=n}\left[\log(-\theta) - \theta\log 2 + \theta\log x_i - \log x_i\right]$$

$$\Rightarrow \log L(\theta) = (\theta-1)\sum_{i=1}^{n}\left[\log x_i - n\theta\log 2\right] + n\log(-\theta)$$

مشتق نسبت به θ و برابر صفر قرار دادیم

$$\left(\log L(\theta)\right)' = \frac{n}{\theta} + \sum_{i=1}^{n}\log x_i - n\log 2 = 0$$

$$\Rightarrow \theta = \frac{-n}{-n\log 2 + \sum_{i=1}^{n}\log x_i}$$

سوال ۳)

الف)

در این مثال، یک تاس غیرمنصفانه داریم که احتمال آمدن عدد ۱ با مقدار p تعیین می‌شود. داده‌های تجربی نشان می‌دهند که در ۱۶ بار پرتاب، ۱۵ بار عدد ۱ و ۱ بار عدد ۰ ظاهر شده است.

فرض کنید که X تعداد دفعاتی باشد که عدد ۱ ظاهر می‌شود و در این حالت X=15 . در اینجا، احتمال آمدن عدد ۱ در هر پرتاب p و احتمال آمدن عدد ۰ (یعنی عدم آمدن ۱) برابر p-1 است.

تابع likelihoodبرای این داده ها به صورت زیر است و در ادامه راه حل داریم:

$$L(p) = (1-p)\cdot p^{15} \xrightarrow[\text{استفاده می‌کنیم}]{\text{log likelihood از}} \ln(L(p)) = \ln(1-p) + 15\ln p$$

مشتق نسبت به p می‌گیریم و برابر صفر قرار می‌دهیم

$$\Longrightarrow \left(\ln(L(p))\right)' = \frac{-1}{1-p} + \frac{15}{p} = 0$$

$$\Rightarrow p = \frac{15}{16}$$

تفسیر :این روش به ما می‌گوید که بر اساس داده‌های مشاهده شده، احتمال آمدن عدد ۱ برابر با ۱۵/۱۶ است. به عبارت دیگر، از آنجا که در ۱۵ بار از ۱۶ بار پرتاب عدد ۱ آمده است، احتمال آمدن عدد ۱ به طور تقریبی ۰.۹۳۷۵ است . .این نتیجه با داده های مشاهده شده سازگار است.

ب)

حالا فرض می کنیم که تعداد پرتاب‌ها ۸ بار باشد و از این ۸ پرتاب، ۷ بار عدد ۱ و ۱ بار عدد ۰ آمده باشد. مشابه به بخش اول داریم:

$$L(p) = (1-p)\times p^{7} \xrightarrow{\text{log likelihood}} \ln(L(p)) = \ln(1-p) + 7\ln p$$

مشتق نسبت به p می‌گیریم برابر صفر قرار می‌دهیم

$$\Longrightarrow \left(\ln(L(p))\right)' = \frac{-1}{1-p} + \frac{7}{p} = 0$$

$$\Rightarrow p = \frac{7}{8}$$

مقایسه دو مجموعه داده:

۱. **برای ۱۶ بار پرتاب :**تخمین p=15/16=0.9375

۲. **برای ۸ بار پرتاب :**تخمین p=7/8=0.875

تأثیر تعداد داده‌ها بر دقت تخمین:

- وقتی تعداد پرتاب‌ها بیشتر باشد، تخمین دقیق‌تر خواهد بود. در این مثال، وقتی ۱۶ بار پرتاب انجام می‌شود، تخمین p به ۰.۹۳۷۵ نزدیک‌تر است. در حالی که وقتی تعداد پرتاب‌ها فقط ۸ بار است، تخمین p به ۰.۸۷۵ می‌رسد.
- این تغییرات به دلیل آن است که با افزایش تعداد مشاهدات، شانس مشاهده نمونه‌هایی که احتمالات واقعی را نمایان کنند، بیشتر می‌شود و بنابراین تخمین دقیق‌تری بدست می‌آید