Chris Piech
CS109

Section #7
Nov 17, 2017

Section #7: Maximum Likelihood Honor Code

1. Single Match:

Let A_i be the event that decision point i is matched. We note that a match occurs when both students make the more popular choice or when both students make the less popular choice. $P(A_i) = P(\text{Both more popular}) + P(\text{Both less popular}) = p^2 + (1-p)^2$.

Let M be a random variable for the number of matches. It is easy to see that each of the 1000 decisions is an independent Bernoulli experiment with probability of success $p' = p^2 + (1-p)^2$. Therefore $M \sim Bin(1000, p')$.

We can use a Normal distribution to approximate a binomial. We approximate $M \sim Bin(1000, p')$ with Normal random variable $Y \sim N(1000p', 1000(1-p')p')$.

2. Maximum Match:

For this problem, we use Maximum Likelihood Estimator (MLE) to estimate the parameters $\theta = (\mu, \beta)$.

$$
\begin{aligned}
L(\theta) &= \prod_{i=1}^{n} f(Y_i \mid \theta) \\
LL(\theta) &= \log \prod_{i=1}^{n} f(Y_i \mid \theta) \\
&= \sum_{i=1}^{n} \log f(Y_i \mid \theta) \\
&= \sum_{i=1}^{n} \log \frac{1}{\beta} e^{-(Z_i + e^{-Z_i})} \qquad \text{where } Z_i = \frac{Y_i - \mu}{\beta} \\
&= \sum_{i=1}^{n} \log \frac{1}{\beta} + \sum_{i=1}^{n} -(Z_i + e^{-Z_i}) \\
&= -n \log(\beta) + \sum_{i=1}^{n} -(Z_i + e^{-Z_i})
\end{aligned}
$$

Now we must choose the values of $\theta = (\mu, \beta)$ that maximize our log-likelihood function. To solve this argmax, we will use Gradient Descent. First, we need to find the first derivative of the log-likelihood function with respect to our parameters.

$$
\begin{aligned}
\frac{\partial LL(\theta)}{\partial \mu} &= \frac{\partial}{\partial \mu}\Big[-n\log(\beta) + \sum_{i=1}^{n} -(Z_i + e^{-Z_i})\Big] \\
&= \sum_{i=1}^{n} \frac{\partial}{\partial \mu}\Big[-(Z_i + e^{-Z_i})\Big] \\
&= \sum_{i=1}^{n} \frac{\partial}{\partial Z_i}\Big[-(Z_i + e^{-Z_i})\Big]\frac{\partial Z_i}{\partial \mu} && \text{By the Chain Rule} \\
&= \sum_{i=1}^{n} \Big[-1 + e^{-Z_i}\Big]\Big[-\frac{1}{\beta}\Big] \\
&= \frac{1}{\beta}\sum_{i=1}^{n} 1 - e^{-Z_i}
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial LL(\theta)}{\partial \beta} &= \frac{\partial}{\partial \beta}\Big[-n\log(\beta) + \sum_{i=1}^{n} -(Z_i + e^{-Z_i})\Big] \\
&= -\frac{n}{\beta} + \sum_{i=1}^{n} \frac{\partial}{\partial \beta}\Big[-(Z_i + e^{-Z_i})\Big] \\
&= -\frac{n}{\beta} + \sum_{i=1}^{n} \frac{\partial}{\partial Z_i}\Big[-(Z_i + e^{-Z_i})\Big]\frac{\partial Z_i}{\partial \beta} && \text{By the Chain Rule} \\
&= -\frac{n}{\beta} + \sum_{i=1}^{n} \Big[-1 + e^{-Z_i}\Big]\Big[\frac{\mu - Y_i}{\beta^2}\Big]
\end{aligned}
$$

3. Understanding:

$P(Y >= 90) = 0.00000017180200395650047$, or nearly 1 in 6 million.

```
from scipy.stats import gumbel_r
print(1 - gumbel_r.cdf(90, 9, 5.2))
```