CS 361 Spring 2018
Homework 8

Nathaniel Murphy (njmurph3)

9.1

(a)

Show that MLE of a dataset following the normal distibution is mean($\{x\}$).

$$\mathcal{L}(\Theta) = P(x_1, \ldots, x_N \mid \Theta,\ \sigma)$$

$$= P(x_1 \mid \Theta,\ \sigma)P(x_2 \mid \Theta,\ \sigma)\ldots P(x_N \mid \Theta,\ \sigma)$$

$$= \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x_i-\Theta)^2}{2\sigma^2}}$$

$$\log\ \mathcal{L}(\Theta) = \sum_{i=1}^{N} \frac{1}{\sqrt{2\pi}\sigma} + \sum_{i=1}^{N} -\frac{(x_i-\Theta)^2}{2\sigma^2}$$

$$= \frac{N}{\sqrt{2\pi}\sigma} - \frac{1}{\sigma^2}\sum_{i=1}^{N}(x_i-\Theta)^2$$

$$\frac{\partial}{\partial\Theta}\log\ \mathcal{L}(\Theta) = 0 - \frac{1}{\sigma^2}\sum_{i=1}^{N}(x_i-\Theta)$$

Set the derivative equal to zero.

$$0 = \frac{1}{\sigma^2}\sum_{i=1}^{N}(x_i-\Theta)$$

$$0 = \sum_{i=1}^{N}(x_i) - N\Theta$$

$$\Theta = \frac{\sum_{i=1}^{N} x_i}{N} = \text{mean}(\{x\})$$

(b)

Show that the MLE for the standard deviation is std($\{x\}$).

$$\mathcal{L}(\Theta) = P(x_1, \ldots, x_N \mid \Theta,\ \sigma)$$

$$= P(x_1 \mid \Theta,\ \mu)P(x_2 \mid \Theta,\ \mu)\ldots P(x_N \mid \Theta,\ \mu)$$

$$= \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}\Theta} e^{\frac{-(x_i-\mu)^2}{2\Theta^2}}$$

$$\log \mathcal{L}(\Theta) = \sum_{i=1}^{N} \log\left(\frac{1}{\sqrt{2\pi}\Theta}\right) + \sum_{i=1}^{N} -\frac{(x_i-\mu)^2}{2\Theta^2}$$

$$= N\log(1) - N\log(\sqrt{2\pi}) - N\log(\Theta) - \frac{1}{\Theta^2}\sum_{i=1}^{N}\frac{(x_i-\mu)^2}{2}$$

$$\frac{\partial}{\partial\Theta}\log \mathcal{L}(\Theta) = -\frac{N}{\Theta} + \frac{2}{\Theta^3}\sum_{i=1}^{N}\frac{(x_i-\mu)^2}{2}$$

Set derivative equal to zero.

$$0 = -\frac{N}{\Theta} + \frac{2}{\Theta^3}\sum_{i=1}^{N}\frac{(x_i-\mu)^2}{2}$$

$$\frac{N}{\Theta} = \frac{2}{\Theta^3}\sum_{i=1}^{N}\frac{(x_i-\mu)^2}{2}$$

$$\Theta^2 = \frac{2}{N}\sum_{i=1}^{N}\frac{(x_i-\mu)^2}{2}$$

$$\Theta = \sqrt{\frac{\sum_{i=1}^{N}(x_i-\mu)^2}{N}} = \text{std}(\{x\})$$

(c)

If all numbers $x_1, \ldots, x_N$ take the same value, we have mean($\{x\}$) $= x_i \ \forall\ i \in [1, N]$.

$$\text{std}(\{x\}) = \sqrt{\frac{\sum_{i=1}^{N}(x_i-x_i)^2}{N}} = \sqrt{0} = 0$$

9.3

MLE for Poisson distribution is $\frac{\sum_{i=1}^{N} n_i}{N}$.

(a)

Day 1: $\frac{(3+1+4+2)}{4} = \frac{10}{4} = 2.5$ pop-ups per hour.
Day 2: $\frac{(2+1+2)}{3} = \frac{5}{3} \approx 1.67$ pop-ups per hour.
Day 3: $\frac{(3+2+2+1+4)}{5} = \frac{12}{5} = 2.4$ pop-ups per hour.

(b)

Day 4 MLE: $\frac{13}{6} \approx 2.1$ pop-ups per hour.

(c)

MLE for all 4 days:

$$\frac{\#\text{ total pop-ups}}{\#\text{ total hours}} = \frac{40}{18} \approx 2.22 \text{ pop-ups per hour}$$

9.4

(a)

$$36 + \frac{36}{r}$$

(b)

Not very reliable because if on the very first spin the ball lands in a 0 slot, you would assume that all the slots are zeroes.

(c)

$$\frac{\left(\frac{36}{r_1} + \frac{36}{r_2} + \frac{36}{r_3}\right)}{3}$$

9.5

MLE for Binomial model is $\Theta = \frac{k}{N}$, where k is number of successes and N is number of draws.

(a)

$$\Theta_{blue} = \frac{0}{1} = 0$$

(b)

$$\Theta_{blue} = \frac{3}{10} = 0.3$$

9.7

(a)

$$\log \frac{P(y=1 \mid x,\ \Theta)}{P(y=0 \mid x,\ \Theta)} = x^T\Theta$$

$$\frac{P(y=1 \mid x,\ \Theta)}{P(y=0 \mid x,\ \Theta} = e^{x^T\Theta}$$

$$P(y=1 \mid x,\ \Theta) = P(y=0 \mid x,\ \Theta)e^{x^T\Theta}$$

$$P(y=1 \mid x,\ \Theta) = (1 - P(y=1 \mid x,\ \Theta))e^{x^T\Theta}$$

$$P(y=1 \mid x,\ \Theta) = e^{x^T\Theta} - P(y=1 \mid x,\ \Theta)e^{x^T\Theta}$$

$$P(y=1 \mid x,\ \Theta) + P(y=1 \mid x,\ \Theta)e^{x^T\Theta} = e^{x^T\Theta}$$

$$P(y=1 \mid x,\ \Theta)(1 + e^{x^T\Theta}) = e^{x^T\Theta}$$

$$P(y=1 \mid x,\ \Theta) = \frac{e^{x^T\Theta}}{1 + e^{x^T\Theta}}$$

(b)

$$\mathcal{L}(\Theta) = \prod_{i \in D} P(D \mid \Theta)$$

$$= \prod_{i \in D} P(y_i = 1 \mid x_i,\ \Theta)^{y_i}(1 - P(y_i = 1 \mid x_i, \Theta))^{1-y_i}$$

$$= \prod_{i \in D} \left(\frac{e^{x_i^T\Theta}}{1 + e^{x_i^T\Theta}} \right)^{y_i} \left(1 - \frac{e^{x_i^T\Theta}}{1 + e^{x_i^T\Theta}} \right)^{1-y_i}$$

$$= \prod_{i \in D} \left(\frac{e^{x_i^T\Theta}}{1 + e^{x_i^T\Theta}} \right)^{y_i} \left(\frac{1}{1 + e^{x_i^T\Theta}} \right)^{1-y_i}$$

$$\log\ \mathcal{L}(\Theta) = \sum_{i \in D} y_i \mathrm{log} \left(\frac{e^{x_i^T\Theta}}{1 + e^{x_i^T\Theta}} \right) + (1 - y_i)\mathrm{log} \left(\frac{1}{1 + e^{x_i^T\Theta}} \right)$$

$$= \sum_{i \in D} y_i \mathrm{log}(e^{x_i^T\Theta}) - y_i \log(1 + e^{x_i^T\Theta}) + \log \left(\frac{1}{1 + e^{x_i^T\Theta}} \right) - y_i \mathrm{log} \left(\frac{1}{1 + e^{x_i^T\Theta}} \right)$$

$$\sum_{i \in D} y_i x_i^T\Theta - y_i \log(1 + e^{x_i^T\Theta}) + \log(1) - \log(1 + e^{x_i^T\Theta}) - y_i \log(1) + y_i \log(1 + e^{x_i^T\Theta})$$

$$= \sum_{i \in D} y_i x_i^T\Theta - \log(1 + e^{x_i^T\Theta})$$

(c)

$$\frac{\partial}{\partial \Theta}\mathcal{L}(\Theta) = y_i x_i^T - \left[x_i^T \cdot^{x_i^T} \cdot \frac{1}{1+e^{x_i^T \Theta}}\right]$$
$$= x_i^T \left[y_i - \frac{e^{x_i^T}}{1+e^{x_i^T \Theta}}\right]$$

Setting this derivative equal to zero and solving for Θ will find the Θ that maximizes $\mathcal{L}(\Theta)$.

9.9

(a)

- $P(n \mid z=0) = 0$
- $P(n \mid z=1) = \frac{1}{37}$
- $P(n \mid z=2) = \frac{2}{38}$
- $P(n \mid z=3) = \frac{3}{39}$

(b)

$P(z=0 \mid observations)$ is only non-zero if observations contains no data points that represent a bet not being payed out. Once a bet is not paid out, we know for a fact that the roulette wheel has more than 0 zero slots.

(c)

z=0:

$$P(z=0 \mid observations) = 0 \text{ because a bet was not payed out } 2 > 0 \text{ times}$$

z=1:

$$P(z=1 \mid observations) = \frac{P(z=1)P(observations \mid z=1)}{P(observations)} = \frac{(0.2)\binom{36}{2}\left(\frac{1}{36}\right)^2\left(\frac{35}{36}\right)^{34}}{P(observations)} = \frac{0.03731}{P(observations)}$$

z=2:

$$P(z=2 \mid observations) = \frac{P(z=2)P(observations \mid z=2)}{P(observations)} = \frac{(0.4)\binom{36}{2}\left(\frac{2}{36}\right)^2\left(\frac{34}{36}\right)^{34}}{P(observations)} = \frac{0.1114}{P(observations)}$$

z=3:

$$P(z=3 \mid observations) = \frac{P(z=3)P(observations \mid z=3)}{P(observations)} = \frac{(0.3)\binom{36}{2}\left(\frac{3}{36}\right)^2\left(\frac{33}{36}\right)^{34}}{P(observations)} = \frac{0.068122}{P(observations)}$$

$$\frac{0 + 0.03731 + 0.1114 + 0.068122}{P(observations)} = 1$$

$$\frac{0.21682}{P(observations)} = 1$$

$$P(observations) = 0.21682$$

$$P(z = 0 \mid observations) = \frac{0}{P(observations)} = 0$$

$$P(z = 1 \mid observations) = \frac{0.03731}{P(observations)} = 0.1721$$

$$P(z = 2 \mid observations) = \frac{0.1114}{P(observations)} = 0.5137$$

$$P(z = 3 \mid observations) = \frac{0.068122}{P(observations)} = 0.3142$$

Posterior for $z = \{0, 1, 2, 3\} = \{0, 0.1721, 0.5137, 0.3142\}$