CSE 312: Foundations of Computing II
**Instructor**: Alex Tsun
**Date**: 08/05/20
**Lecture Topics**: 7.3 Method of Moments Estimation, 7.4 Beta/Dirichlet Distributions

[**Tags:** Estimation]

1. Suppose $x = (x_1, \dots, x_n)$ are iid samples from the following distributions. Estimate the parameter(s) using your favorite technique (MLE or MoM). **Hint**: Use MoM.
   a. The $Gamma(r, \lambda)$ distribution. Estimate both $r$ and $\lambda$.
   b. The $Rayleigh(\sigma)$ distribution with density $f_X(x;\sigma) = \frac{x}{\sigma^2} e^{-x^2/2\sigma^2}, x \geq 0$ with expectation $\sigma\sqrt{\frac{\pi}{2}}$.

**Solution**:

a. We have $k = 2$ parameters to estimate. Let $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ and $\overline{x^2} = \frac{1}{n}\sum_{i=1}^{n} x_i^2$ We set

$$E[X] = \bar{x}$$
$$E[X^2] = \overline{x^2}$$

For a $Gamma(r, \lambda)$ rv,

$$E[X] = \frac{r}{\lambda}, \qquad E[X^2] = Var(X) + E[X]^2 = \frac{r}{\lambda^2} + \left(\frac{r}{\lambda}\right)^2 = \frac{r(r+1)}{\lambda^2}$$

So we must solve the two equations:

$$\frac{r}{\lambda} = E[X] = \bar{x}, \qquad \frac{r(r+1)}{\lambda^2} = E[X^2] = \overline{x^2}$$

Let's divide the second equation by the first (note $\bar{x}^2 \neq \overline{x^2}$), then subtract $\frac{r}{\lambda} = \bar{x}$:

$$\frac{\overline{x^2}}{\bar{x}} = \frac{E[X^2]}{E[X]} = \frac{\frac{r(r+1)}{\lambda^2}}{\frac{r}{\lambda}} = \frac{r+1}{\lambda} \quad \rightarrow \quad \frac{1}{\lambda} = \frac{r+1}{\lambda} - \frac{r}{\lambda} = \frac{\overline{x^2}}{\bar{x}} - \bar{x} = \frac{\overline{x^2} - \bar{x}^2}{\bar{x}}$$

Hence, taking the reciprocal gives

$$\hat{\lambda} = \frac{\bar{x}}{\overline{x^2} - \bar{x}^2}$$

Then, since $\frac{r}{\lambda} = \bar{x}$ or equivalently $r = \lambda\bar{x}$, we get

$$\hat{r} = \hat{\lambda}\bar{x} = \frac{\bar{x}^2}{\overline{x^2} - \bar{x}^2}$$

b. We set

$$\sigma\sqrt{\frac{\pi}{2}} = E[X] = \bar{x} \quad \rightarrow \quad \hat{\sigma} = \bar{x}\sqrt{\frac{2}{\pi}}$$

[**Tags:** Beta/Dirichlet]

2. Suppose we roll a (possibly unfair) 4-sided die 29 times. Then, the number of times each digit appears is $\boldsymbol{X} = (X_1, X_2, X_3, X_4) \sim Mult_4(n = 29, \boldsymbol{p})$, where $\boldsymbol{p} = (p_1, p_2, p_3, p_4)$ is unknown. We happened to observe 5 ones, 7 twos, 6 threes, and 11 fours.
    a. A $Beta(\alpha_1, \beta_1)$ rv would be suitable to model our belief on $p_1$ (the probability of rolling a one) with what parameters $\alpha_1, \beta_1$?
    b. A $Beta(\alpha_2, \beta_2)$ rv would be suitable to model our belief on $p_2$ (the probability of rolling a two) with what parameters $\alpha_2, \beta_2$?
    c. Let's instead say we wanted to jointly model all the unknown parameters $\boldsymbol{p}$. A $Dirichlet(\boldsymbol{\gamma})$ would be suitable, more efficient than modelling all four separately, and also enforce that $\sum_{i=1}^{4} p_i = 1$. Which parameter vector $\gamma = (\gamma_1, \gamma_2, \gamma_3, \gamma_4)$ would best model our belief?

**Solution**:

a. We have $Beta(\alpha_1 = 6, \beta_1 = 25)$ since we saw $\alpha_1 - 1 = 5$ ones and $\beta_1 - 1 = 24 = 7 + 6 + 11$ other values.
b. We have $Beta(\alpha_2 = 8, \beta_2 = 23)$ since we saw $\alpha_2 - 1 = 7$ twos and $\beta_2 - 1 = 22 = 5 + 6 + 11$ other values.
c. We have $Dirichlet(\gamma_1 = 6, \gamma_2 = 8, \gamma_3 = 7, \gamma_4 = 12)$ since we saw $\gamma_1 - 1 = 5$ ones, $\gamma_2 - 1 = 7$ twos, $\gamma_3 - 1 = 6$ threes, and $\gamma_4 - 1 = 12$ fours.