

[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, λ) distribution. Estimate both r and λ .
 - b. The *Rayleigh*(σ) 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

$$\begin{aligned} E[X] &= \bar{x} \\ E[X^2] &= \overline{x^2} \end{aligned}$$

For a *Gamma*(r, λ) rv,

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

So we must solve the two equations:

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

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

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

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

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

[Tags: Beta/Dirichlet]

2. Suppose we roll a (possibly unfair) 4-sided die 29 times. Then, the number of times each digit appears is $\mathbf{X} = (X_1, X_2, X_3, X_4) \sim \text{Mult}_4(n = 29, \mathbf{p})$, where $\mathbf{p} = (p_1, p_2, p_3, p_4)$ is unknown. We happened to observe 5 ones, 7 twos, 6 threes, and 11 fours.
- A $\text{Beta}(\alpha_1, \beta_1)$ rv would be suitable to model our belief on p_1 (the probability of rolling a one) with what parameters α_1, β_1 ?
 - A $\text{Beta}(\alpha_2, \beta_2)$ rv would be suitable to model our belief on p_2 (the probability of rolling a two) with what parameters α_2, β_2 ?
 - Let's instead say we wanted to jointly model all the unknown parameters \mathbf{p} . A $\text{Dirichlet}(\boldsymbol{\gamma})$ would be suitable, more efficient than modelling all four separately, and also enforce that $\sum_{i=1}^4 p_i = 1$. Which parameter vector $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \gamma_3, \gamma_4)$ would best model our belief?

Solution:

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