Special Cases 35M

Matching Moments

If we have **complete data**, the MLE of parameters for several distributions can be calculated by matching the fitted raw moments to the empirical raw moments as a shortcut. This means we can calculate the MLE of parameters by setting

$$\mathrm{E}\left[X^k\right] = \frac{\sum_{i=1}^{n} x_i^k}{n}$$

where $\mathrm{E}\left[X^k\right]$ is the fitted k^{th} raw moment and $\frac{\sum_{i=1}^{n} x_i^k}{n}$ is the empirical k^{th} raw moment.

Gamma

For $X \sim \text{Gamma}\,(\alpha,\, \theta)$ where α is fixed, the MLE of θ is the sample mean divided by α.

$$\mathrm{E}[X] = \alpha\hat{\theta} = \bar{x} \qquad \Rightarrow \qquad \hat{\theta} = \frac{\bar{x}}{\alpha}$$

The **exponential** distribution is a special case of the gamma distribution with $\alpha = 1$. Thus, the MLE of θ is the sample mean.

Normal

The normal distribution is the only distribution where we can match the **first two moments** to estimate the two parameters. For $X \sim \text{Normal}\,\left(\mu,\, \sigma^2\right)$,

$$\mathrm{E}[X] = \hat{\mu} = \bar{x}$$

$$\mathrm{E}\left[X^2\right] = \hat{\sigma}^2 + \hat{\mu}^2 = \frac{\sum_{i=1}^n x_i^2}{n}$$

Thus,

$$\begin{aligned}\hat{\sigma}^2 &= \frac{\sum_{i=1}^n x_i^2}{n} - \bar{x}^2 \\ &= \frac{\sum_{i=1}^n \left(x_i - \bar{x}\right)^2}{n}\end{aligned}$$

Coach's Remarks

Keep in mind that we can only estimate σ^2 by matching the 2nd moment when μ is also estimated using MLE.

If μ is fixed, the MLE of σ^2 is

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n \left(x_i - \mu\right)^2}{n}$$

which does **not** equal $\frac{\sum_{i=1}^n x_i^2}{n} - \mu^2$.

When asked to estimate the **lognormal** parameters, we can convert the lognormal random variables to normal and apply the normal shortcuts. For $X \sim \text{Lognormal}\left(\mu,\, \sigma^2\right)$,

$$\hat{\mu} = \frac{\sum_{i=1}^{n} \ln x_i}{n}$$

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n} \left(\ln x_i - \hat{\mu}\right)^2}{n}$$

(a, b, 0) Class

POISSON

For $X \sim \text{Poisson}\ (\lambda)$, the MLE of λ is the sample mean.

$$\text{E}[X] = \hat{\lambda} = \bar{x}$$

BINOMIAL

For $X \sim \text{Binomial}\ (m,\ q)$ where m is fixed, the MLE of q is the sample mean divided by m.

$$\text{E}[X] = m\hat{q} = \bar{x} \qquad \Rightarrow \qquad \hat{q} = \frac{\bar{x}}{m}$$

NEGATIVE BINOMIAL

For $X \sim \text{Negative Binomial}\ (r,\ \beta)$ where r is fixed, the MLE of β is the sample mean divided by r.

$$\mathrm{E}[X] = r\hat{\beta} = \bar{x} \qquad \Rightarrow \qquad \hat{\beta} = \frac{\bar{x}}{r}$$

The **geometric** distribution is a special case of the negative binomial distribution with $r = 1$. Thus, the MLE of β is the sample mean.

The derivations of the matching moment shortcuts are included in the appendix at the end of this section.

(a, b, 1) Class

ZERO-TRUNCATED

The shortcuts above also work for the $(a,\ b,\ 1)$ counterparts. Just make sure to match the appropriate means.

$$\mathrm{E}\left[X^T\right] = \frac{1}{1 - p_0} \cdot \mathrm{E}[X] = \bar{x}$$

where X is a $(a,\ b,\ 0)$ class random variable and X^T is its zero-truncated counterpart.

For **zero-truncated Poisson**,

$$\frac{\hat{\lambda}}{1 - e^{-\hat{\lambda}}} = \bar{x}$$

For **zero-truncated binomial**,

$$\frac{m\hat{q}}{1 - \left(1 - \hat{q}\right)^m} = \bar{x}$$

For **zero-truncated negative binomial**,

$$\frac{r\hat{\beta}}{1-\left(1+\hat{\beta}\right)^{-r}}=\bar{x}$$

ZERO-MODIFIED

To estimate the parameters of zero-modified distributions using the maximum likelihood method, perform these two matches:

- Estimate the fitted probability of zero as the proportion of zeros in the sample.

$$\hat{p}_0^M=\frac{n_0}{n}$$

  where n_0 is the number of observed zeros and n is the sample size.

- Match the fitted mean to the sample mean.

$$\mathbf{E}\left[X^M\right]=\frac{1-\hat{p}_0^M}{1-p_0}\cdot\mathbf{E}[X]=\bar{x}$$

  where X is a $(a,\,b,\,0)$ class random variable and X^M is its zero-modified counterpart.

Let's revisit a couple of examples.

In the first example we had at the beginning of Section S3.1.1:

The annual number of claims is modeled to have a Poisson distribution with mean λ. You observed 3, 1, and 2 claims in the past three years, respectively. Using the data, what is an appropriate estimate of λ?

For a Poisson distribution, the MLE of λ is the sample mean.

$$\begin{aligned}\hat{\lambda} &= \bar{x} \\ &= \frac{3+1+2}{3} \\ &= \mathbf{2}\end{aligned}$$

Example S3.1.1.1 can be solved using the exponential shortcut, as demonstrated below.

Example S3.1.3.1

You observe 4 claims:

$$20 \quad 50 \quad 150 \quad 380$$

The claim amounts follow an exponential distribution with mean θ.

Calculate the maximum likelihood estimate of θ.

Solution

For an exponential distribution, the MLE of θ is the sample mean.

$$\begin{aligned}\hat{\theta} &= \bar{x} \\ &= \frac{20+50+150+380}{4} \\ &= \mathbf{150}\end{aligned}$$

Example S3.1.1.2 can be solved using the gamma shortcut, as demonstrated below.

Example S3.1.3.2

Losses follow a distribution with probability density function:

$$f(x) = \frac{x^3 e^{-x/\theta}}{6\theta^4}, \quad x > 0$$

You observe 5 losses totaling 700.

Calculate the maximum likelihood estimate of θ.

Solution

The PDF has a term x raised to a positive power, multiplied by e raised to a negative multiple of x. It is a gamma PDF.

$$f(x) = \frac{x^3 e^{-x/\theta}}{6\theta^4} = c \cdot x^3 e^{-x/\theta}$$

Compare it to the generic gamma PDF provided by the exam table:

$$\frac{(x\,/\,\theta)^{\alpha}e^{-x\,/\,\theta}}{x\cdot\Gamma(\alpha)}$$

Conclude that losses follow a gamma distribution with parameters $\alpha = 4$ and θ.

For a gamma distribution with fixed α, the fitted mean equals the sample mean.

$$\begin{aligned}\alpha\hat{\theta} &= \bar{x}\\ 4\hat{\theta} &= \frac{700}{5}\\ \hat{\theta} &= \frac{140}{4} = \mathbf{35}\end{aligned}$$

■

Example S3.1.1.3 shows the derivation of the normal shortcuts. Let's try a different example that applies normal shortcuts to a lognormal case.

Example S3.1.3.3

You observe the following claims from a dataset:

$$25 \qquad 70 \qquad 215 \qquad 535$$

You fit a lognormal distribution to the data using maximum likelihood estimation.

Determine the mean of the fitted distribution.

Solution

Let X represent the claim variable, so $X \sim \text{Lognormal}\left(\mu,\, \sigma^2\right)$. Then,

$$\ln X \sim \text{Normal}\left(\mu,\, \sigma^2\right)$$

We can estimate μ and σ^2 using the normal shortcuts.

$$\begin{aligned}\hat{\mu} &= \frac{\sum_{i=1}^{4} \ln x_i}{4} \\ &= \frac{\ln 25 + \ln 70 + \ln 215 + \ln 535}{4} \\ &= 4.7801\end{aligned}$$

$$\begin{aligned}\hat{\sigma}^2 &= \frac{\sum_{i=1}^{4} \left(\ln x_i\right)^2}{4} - \hat{\mu}^2 \\ &= \frac{\left(\ln 25\right)^2 + \left(\ln 70\right)^2 + \left(\ln 215\right)^2 + \left(\ln 535\right)^2}{4} - 4.7801^2 \\ &= 1.3313\end{aligned}$$

Look up the lognormal distribution's mean formula in the exam table.

$$\begin{aligned}\hat{\mathrm{E}}[X] &= e^{\hat{\mu} + 0.5\hat{\sigma}^2} \\ &= e^{4.7801 + 0.5(1.3313)} \\ &= \mathbf{231.7657}\end{aligned}$$

Example S3.1.3.4

You are given the following accident data from 100 insurance policies:

Number of Accidents	Number of Policies
0	45
1	37
2	16
3	2
4+	0

You fit a zero-modified binomial distribution with $m = 3$ to the data using the maximum likelihood technique.

Determine the second raw moment of the fitted distribution, $\mathrm{E}\left[\left(N^M\right)^2\right]$.

Solution

First, estimate the fitted probability of zero as the sample's proportion of zeros.

$$\hat{p}_0^M = \frac{45}{100} = 0.45$$

Second, match the fitted mean to the sample mean.

$$\begin{aligned}
\frac{1-\hat{p}_0^M}{1-p_0}\cdot \mathrm{E}[N] &= \bar{x} \\
\frac{1-0.45}{1-(1-q)^3}\cdot 3q &= \frac{45(0)+37(1)+16(2)+2(3)}{100} \\
\frac{1.65q}{1-(1-3q+3q^2-q^3)} &= 0.75 \\
\frac{q}{3q-3q^2+q^3} &= 0.4545 \\
\frac{1}{0.4545} &= 3-3q+q^2 \\
q^2-3q+0.8 &= 0
\end{aligned}$$

Using the quadratic formula,

$$\begin{aligned}
\hat{q} &= \frac{-(-3)\pm\sqrt{(-3)^2-4(1)(0.8)}}{2(1)} \\
&= 2.7042 \quad \text{or} \quad 0.2958
\end{aligned}$$

q needs to be between 0 and 1. Therefore, the MLE of q is 0.2958.

The zero-modified second moment can be calculated by multiplying the original second moment by the same factor.

$$\begin{aligned}
\mathrm{E}\left[\left(N^M\right)^2\right] &= \frac{1-p_0^M}{1-p_0}\cdot \mathrm{E}\left[N^2\right] \\
&= \frac{1-p_0^M}{1-(1-q)^3}\left[3q(1-q)+(3q)^2\right]
\end{aligned}$$

Substitute the estimated p_0^M and q to estimate the second moment.

$$\begin{aligned}
\hat{\mathrm{E}}\left[\left(N^M\right)^2\right] &= \frac{1-0.45}{1-(1-0.2958)^3}\left[3(0.2958)(1-0.2958)+[3(0.2958)\right. \\
&= \mathbf{1.1938}
\end{aligned}$$

Uniform

For a uniform distribution on the interval $[0,\ \theta]$, the MLE of θ is the largest observation.

Consider the following example:

You are given the following observations:

$$2 \qquad 5 \qquad 8 \qquad 13 \qquad 16$$

You fit a uniform distribution on $[0,\ \theta]$ to this data using the maximum likelihood method.

Determine $\hat{\theta}$.

The PDF of the uniform distribution is a constant regardless of the observation.

$$f(x) = \frac{1}{\theta}$$

Using first principles,

$$L(\theta) = \frac{1}{\theta^5}$$

$$l(\theta) = -5 \ln \theta$$

$$l'(\theta) = -\frac{5}{\theta}$$

We will end up with this equation, which does not have a solution.

$$-\frac{5}{\theta} = 0$$

Thus, we have to think of another way to determine θ.

Recall that our goal is to maximize the likelihood function. The likelihood function increases as θ decreases.

$$\theta \downarrow \qquad \rightarrow \qquad L(\theta) = \frac{1}{\theta^5} \uparrow$$

Thus, we want θ as small as possible.

Meanwhile, θ is the upper limit of the distribution's range. Thus, it needs to be at least the greatest observed value.

$$\theta \geq \max(x_1, x_2, \ldots, x_n)$$

Based on the two requirements above, the MLE of θ is the greatest observed value.

$$\begin{aligned}\hat{\theta} &= \max(2,\ 5,\ 8,\ 13,\ 16) \\ &= \mathbf{16}\end{aligned}$$

Coach's Remarks

Keep in mind that the shortcuts above are only applicable if there is **complete data**, i.e. no truncation, censoring and grouping.

Other Shortcuts

Besides the shortcuts discussed above, there are several MLE estimates that can be derived in a closed form expression. They are summarized in the table below.

In this table,

- n is the number of uncensored data points;
- c is the number of censored data points;
- x_i is the observed value, or the censoring point for censored data;
- d_i is the truncation point.

Distribution	Shortcut
Pareto, fixed θ	$\hat{\alpha} = \dfrac{n}{\sum_{i=1}^{n+c} [\ln(x_i + \theta) - \ln(d_i + \theta)]}$
S-P Pareto, fixed θ	$\hat{\alpha} = \dfrac{n}{\sum_{i=1}^{n+c} \{\ln x_i - \ln[\max(\theta,\ d_i)]\}}$
Exponential	$\hat{\theta} = \dfrac{\sum_{i=1}^{n+c} (x_i - d_i)}{n}$
Inverse exponential	$\hat{\theta} = \dfrac{n}{\sum_{i=1}^{n} 1/x_i}$
Weibull, fixed τ	$\hat{\theta} = \left(\dfrac{\sum_{i=1}^{n+c} x_i^\tau - \sum_{i=1}^{n+c} d_i^\tau}{n}\right)^{1/\tau}$

Distribution	Shortcut
Beta, fixed $\theta, b = 1$	$\hat{a} = \dfrac{n}{n \ln \theta - \sum_{i=1}^{n} \ln x_i}$
Beta, fixed $\theta, a = 1$	$\hat{b} = \dfrac{n}{n \ln \theta - \sum_{i=1}^{n} \ln (\theta - x_i)}$
Uniform $[0,\ \theta]$, grouped data	$\hat{\theta} = c_j \left(\dfrac{n}{n - n_j} \right)$ n_j: number of data points in the last interval $[c_j,\ \infty)$

From the table, the following shortcuts can be applied only when **complete data** is available:

- Inverse exponential
- Beta, fixed $\theta, b = 1$
- Beta, fixed $\theta, a = 1$

Also, the uniform shortcut cannot be used on truncated data

Coach's Remarks

We do not recommend memorizing all the shortcuts in the table. We included them for the sake of completeness. The benefit of these shortcuts is marginal, since it would likely impact just a few problems on the exam. Either way, you are expected to practice solving MLE problems using first principles. Having said that, you may wish to learn the exponential shortcut, which we'll demonstrate in the next example.

The derivations of the shortcuts are included in the appendix at the end of this section.

Example S3.1.3.5

An insurance company offers two types of policies:

- Policy I has no deductible and a policy limit of 100.
- Policy II has a deductible of 20 and no policy limit.

You are given the following samples from these two policies. Losses below the deductible are not recorded.

Policy	Observed Losses						
I	50	60	60	70	80	100^+	100^+
II	30	50	60	70	90	120	

$+$ indicates that the loss exceeded the policy limit.

An actuary fits a ground-up exponential distribution using maximum likelihood estimation.

Estimate the mean of the ground-up distribution.

Solution

Let x_i be the i^{th} observed loss, and d_i be the deductible that x_i is subject to. n and c are the numbers of uncensored and censored losses, respectively. The MLE of θ is

$$\hat{\theta} = \frac{\sum_{i=1}^{n+c} (x_i - d_i)}{n}$$

For Policy I, there are 5 uncensored observations.

x_i	d_i	$x_i - d_i$
50	0	50
60	0	60
60	0	60
70	0	70
80	0	80
100	0	100
100	0	100

For Policy II, there are 6 uncensored observations.

x_i	d_i	$x_i - d_i$
30	20	10
50	20	30
60	20	40
70	20	50
90	20	70
120	20	100

Use the exponential shortcut to calculate the MLE of θ.

$$\begin{aligned}\hat{\theta} &= \frac{(50+60+\ldots+100)+(10+30+\ldots+100)}{5+6} \\ &= \frac{820}{11} \\ &= \mathbf{74.5455}\end{aligned}$$

Coach's Remarks

The shortcut can be memorized as