Answers to Exercise Problems

Q6.1 [Ans $p \approx 0.0062$]

Q6.2 [Ans: $(a)p \approx 0.9192$, (b) $p \approx 0.9973$]

Q6.3 [Ans: (a) $p \approx 0.4286$; (b) 24.63 hr; (c) 828.82 hr; 1487.42 hr]

Q6.4 [Ans: (a) $p \approx 0.4213$; (b) 1743 hr]

Q6.5 [Ans: (a) 10^4 hr; (b) 6931.47 hr; (c) $p \approx 0.6321$]

Q6.6 [Ans: (a) $p \approx 0.2466$; (b) $p \approx 0.2466$; (c) 1151.3 hr]

Q6.7 [Ans: (b) 2×10^4 hr; (c) 1.99×10^4 hr]

Q6.8 [Ans: (a) 10000 hr; (c) $p \approx 0.1353$]

Q6.9 [Ans: 200 hr; 1.274]

Q6.10 [Ans: (a) $p \approx 0.7836$; (b) 514.54 hr]

$\boxed{7}$ Distribution Selection

Method of Moments (MoM)

MoM is a technique for estimating parameters of probability distributions that is based on *matching* the *sample moments* with the corresponding *distribution moments*.

Distribution moments:

1st (raw) moment (mean): $\mu = \int_{-\infty}^{\infty} x f(x) dx$

2nd (central) moment (variance): $\mu_2 = \int_{-\infty}^{\infty} (x-\mu)^2 f(x) dx$

3rd (central) moment: $\mu_3 = \int_{-\infty}^{\infty} (x-\mu)^3 f(x) dx$

kth (central) moment: $\mu_k = \int_{-\infty}^{\infty} (x-\mu)^k f(x) dx$

Method of Moments (MoM) (cont'd)

Sample moments:

Sample mean: $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$

Sample variance: $s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$

Sample 3rd (central) moment: $m_3 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^3}{n}$

Sample kth (central) moment: $m_k = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^k}{n}$

MoM for Exponential Distribution

– Exponential PDF:

$$f(x) = \lambda e^{-\lambda x}, \quad x \geq 0$$

– Distribution parameter: λ

– Relation with sample mean: Dist. mean = sample mean

$$\frac{1}{\lambda} = \bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

MoM for Normal Distribution

– Normal PDF:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right], \quad -\infty < x < \infty$$

– Distribution parameters: μ, σ

– Relation with sample estimates: Dist. mean/SD = sample mean/SD

$$\mu = \bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

$$\sigma^2 = s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$$

MoM for Lognormal Distribution

– Lognormal PDF:

$$f(x) = \frac{1}{x\omega\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\ln x - \theta}{\omega}\right)^2\right], \quad 0 < x < \infty$$

– Distribution parameters: θ, ω

– Relation with sample estimates: Dist. mean/SD = sample mean/SD

$$\mu = e^{\theta + \omega^2/2} = \frac{\sum_{i=1}^{n} x_i}{n} = \bar{x}$$

$$\sigma^2 = e^{2\theta + \omega^2}(e^{\omega^2} - 1) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} = s^2$$

MoM for Gamma Distribution

– Gamma PDF:

$$f(x) = \frac{\lambda(\lambda x)^{r-1} e^{-\lambda x}}{\Gamma(r)}, \quad x \geq 0$$

– Distribution parameters: r, λ

– Relation with sample estimates: Dist. mean/SD = sample mean/SD

$$\mu = \frac{r}{\lambda} = \frac{\sum_{i=1}^{n} x_i}{n} = \bar{x}$$

$$\sigma^2 = \frac{r}{\lambda^2} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} = s^2$$

MoM for Weibull Distribution

- Weibull PDF:

$$f(x) = \frac{\beta}{\theta}\left(\frac{x}{\theta}\right)^{\beta-1} \exp\left[-\left(\frac{x}{\theta}\right)^{\beta}\right], \quad x \geq 0$$

- Distribution parameters: θ, β

- Relation with sample estimates: Dist. mean/SD = sample mean/SD

$$\mu = \theta\Gamma\left(1+\frac{1}{\beta}\right) = \frac{\sum_{i=1}^{n} x_i}{n} = \bar{x}$$

$$\sigma^2 = \theta^2\left[\Gamma\left(1+\frac{2}{\beta}\right) - \left\{\Gamma\left(1+\frac{1}{\beta}\right)\right\}^2\right] = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1} = s^2$$

MoM Example

- Data on aluminium contamination (ppm) in plastic:

Aluminum Contamination (ppm)

30	30	60	63	70	79	87
90	101	102	115	118	119	119
120	125	140	145	172	182	
183	191	222	244	291	511	

From "The Lognormal Distribution for Modeling Quality Data When the Mean Is Near Zero," *Journal of Quality Technology*, 1990, pp. 105–110.

- Sample mean, $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = 142.6538$ ppm, where $n = 26$.

- Sample variance, $s^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1} = 98.2043^2$.

MoM Example (cont'd)

- Let us assume that the *data follows the exponential* distribution.

- Exponential PDF $(x \geq 0)$:

$$f(x) = \lambda e^{-\lambda x}$$

- Distribution parameter: λ

- Relation with sample estimates:

$$\frac{1}{\lambda} = \bar{x} = 142.6538 \implies \lambda = 0.00701$$

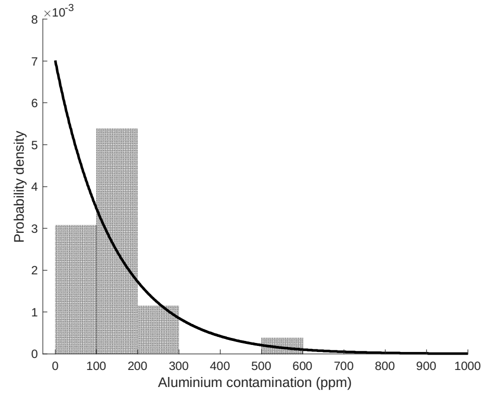


Figure: **Exponential distribution** fitted to the aluminium contamination data using method of moments.

MoM Example (cont'd)

- Let us assume that the *data follows the normal* distribution.

- Normal PDF $(-\infty < x < \infty)$:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$$

- Distribution parameters: μ, σ

- Relation with sample estimates:

$$\mu = \bar{x} = 142.6538$$

$$\sigma = \sqrt{s^2} = 98.2043$$

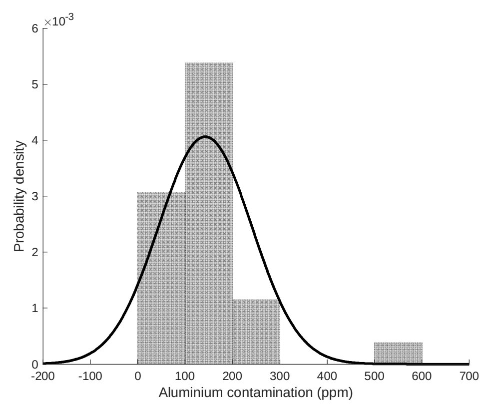


Figure: **Normal distribution** fitted to the aluminium contamination data using method of moments.

MoM Example (cont'd)

– Let us assume that the *data follows the lognormal* distribution.

– Lognormal PDF $(0 < x < \infty)$:

$$f(x) = \frac{1}{x\omega\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\ln x - \theta}{\omega}\right)^2\right]$$

– Distribution parameters: θ, ω

– Relation with sample estimates:

$$\omega = \sqrt{\ln\left(1 + \frac{s^2}{\bar{x}^2}\right)} = 0.6228$$

$$\theta = \ln \bar{x} - \frac{\omega^2}{2} = 4.7665$$

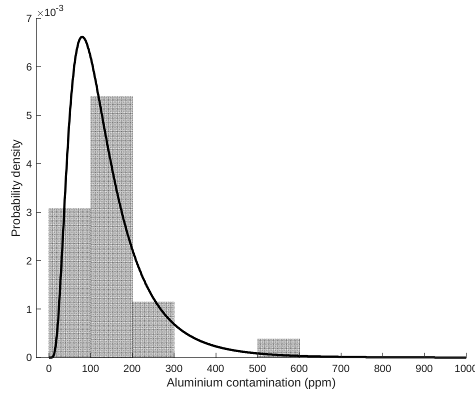


Figure: **Lognormal distribution** fitted to the aluminium contamination data using method of moments.

MoM Example (cont'd)

– Let us assume that the *data follows the gamma* distribution.

– Gamma PDF $(x \geq 0)$:

$$f(x) = \frac{\lambda(\lambda x)^{r-1} e^{-\lambda x}}{\Gamma(r)}$$

– Distribution parameters: r, λ

– Relation with sample estimates:

$$r = \frac{\bar{x}^2}{s^2} = 2.1101, \qquad \lambda = \frac{\bar{x}}{s^2} = 0.0148$$

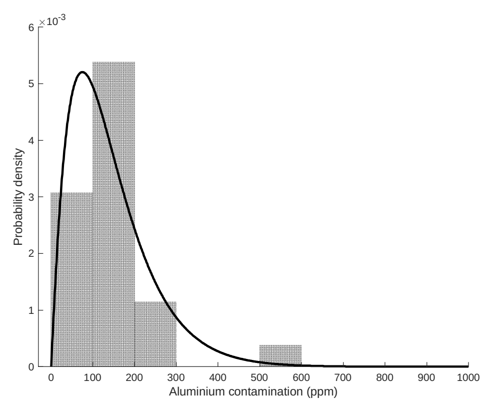


Figure: **Gamma distribution** fitted to the aluminium contamination data using method of moments.

MoM Example (cont'd)

– Let us assume that the *data follows the Weibull* distribution.

– Weibull PDF $(x \geq 0)$:

$$f(x) = \frac{\beta}{\theta}\left(\frac{x}{\theta}\right)^{\beta-1} \exp\left[-\left(\frac{x}{\theta}\right)^{\beta}\right]$$

– Distribution parameters: θ, β

– Relation with sample estimates:

$$\mu = \theta\Gamma\left(1+\tfrac{1}{\beta}\right) = \bar{x} = 142.6538$$

$$\sigma^2 = \theta^2\left[\Gamma\left(1+\tfrac{2}{\beta}\right) - \left\{\Gamma\left(1+\tfrac{1}{\beta}\right)\right\}^2\right] = s^2 = 98.2043^2$$

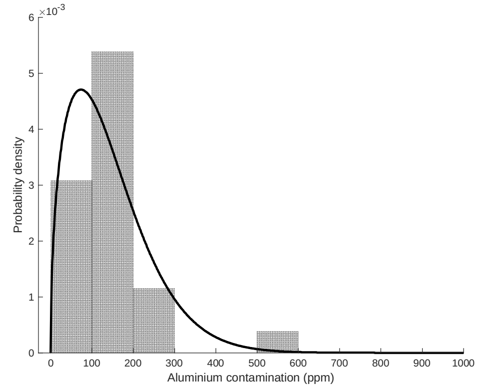


Figure: **Weibull distribution** fitted to the aluminium contamination data using method of moments (**solved numerically**: $\theta = 157.727$, $\beta = 1.478$).

MoM Example (cont'd)

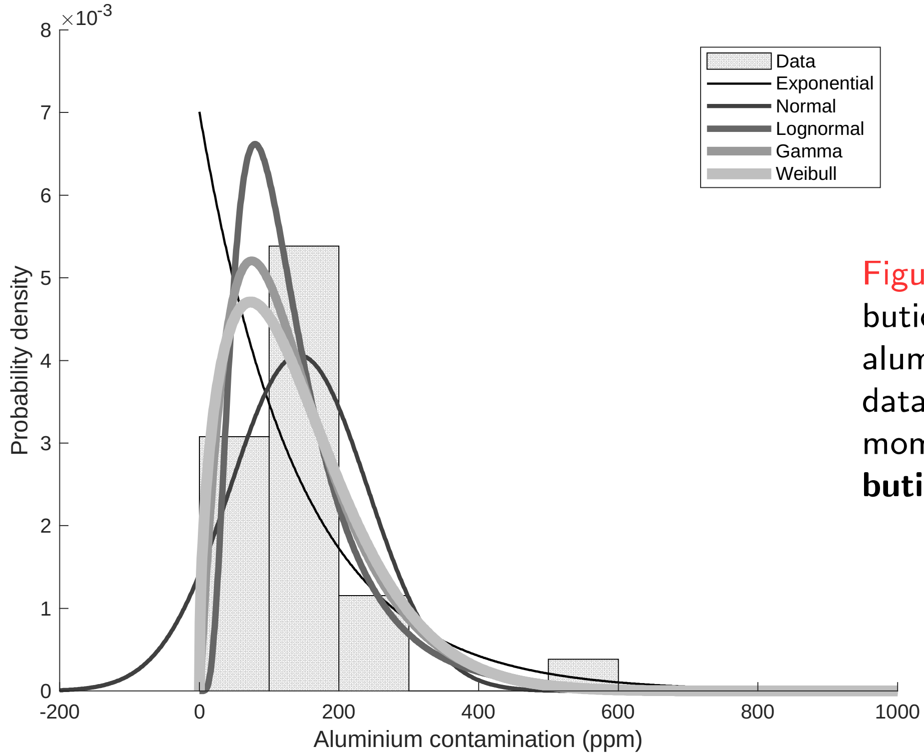


Figure: Various distributions fitted to the aluminium contamination data using method of moments. **Which distribution fits well**?

Probability Paper Plots (PPP)

PPP is a graphical approach used in statistics to assess whether a dataset follows a particular probability distribution. The *scale of the graph paper* should be such that a *linear relationship* is produced for the plotted data.

- The objective is to write the cumulative distribution function (CDF) in **linear form**: $y = mx + c$. The distribution parameters can be estimated from the slope m and intercept c.
- The linearity or the **lack of linearity** can then be used as a basis for determining whether a sample data have come from a particular distribution.
- Statistical tests can then be used to quantitatively assess the goodness of fit between the data and the chosen distribution.

Steps for Constructing Probability Paper Plots

- In order to plot the sample data against a particular probability density function, one need's to **assign a probability** associated with each data point.
- **Arrange the data** in an increasing order: $x_1, x_2, \cdots, x_n$ $(x_{i+1} > x_i)$, and calculate the **rank probability** associated with each data point (mean rank assumption) as:

$$P_i = \frac{i}{n+1}$$

  where i $(i = 1, 2, \cdots, n)$ is the rank of the data and n is the number of data points.
- We then **equate** the rank probability of the data with the chosen cumulative distribution function (CDF) as:

$$P_i = F(x_i) \implies \frac{i}{n+1} = F(x_i) \implies F^{-1}\left(\frac{i}{n+1}\right) = x_i$$

  where $F^{-1}(\cdot)$ is the inverse CDF of the chosen distribution.

Constructing Normal Probability Paper Plots

- Equate the rank probability of the data with the normal CDF:

$$P_i = F(x_i) = \Phi\left(\frac{x_i - \mu}{\sigma}\right)$$
$$\implies \Phi^{-1}\left(\frac{i}{n+1}\right) = \frac{x_i - \mu}{\sigma} \quad (\text{replace: } P_i = \tfrac{i}{n+1})$$
$$\implies x_i = \sigma\Phi^{-1}\left(\frac{i}{n+1}\right) + \mu \quad (\text{form: } y = mx + c)$$

  where $\Phi(\cdot)$ is the standard normal CDF; $\Phi^{-1}(\cdot)$ is the standard normal inverse CDF.
- The normal PPP is constructed by plotting the data x_i against the corresponding percentile value given by $\Phi^{-1}(P_i)$.
- Distribution parameters can be estimated from the above straight line equation as: mean (μ) = intercept(c) and standard deviation (σ) = slope(m).

Constructing Normal Probability Paper Plots

- Once x_i values are plotted against $\Phi^{-1}(P_i)$ values, they should form a straight line (approximately) on a normal PPP if the data belongs to a normal distribution.
- The best-fit line is obtained through **linear regression** which minimizes the sum of squares of deviations between the data and the fitted line.
- The **coefficient of determination** R^2 (normally ranges from 0 to 1) gives information about the goodness of fit of the model:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(x_i - f_i)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

  where $f_i = m\Phi^{-1}(P_i) + c$ is predicted from the best-fit line and $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$.
- The R^2 value reflects the strength of the linear association of the data. The **higher** the R^2 value, the **better the fit**.

Example: Normal PPP

– A total of 25 samples of concrete strength were taken during the construction of a building and tested for crushing strength in the lab. The data set is shown below.

Test No.	Crushing Strength (MPa)	Test No.	Crushing Strength (MPa)	Test No.	Crushing Strength (MPa)
1	40.0	10	33.8	19	44.1
2	37.9	11	39.3	20	40.0
3	29.0	12	42.1	21	40.7
4	31.7	13	45.5	22	42.1
5	39.3	14	41.4	23	38.6
6	40.7	15	47.6	24	36.5
7	42.7	16	38.6	25	40.7
8	40.0	17	35.9		
9	50.3	18	41.4		

Example: Normal PPP (cont'd)

– *Step 1:* Arrange (or sort) the data in an increasing order.

– *Step 2:* Calculate the probability based on the rank assumption.

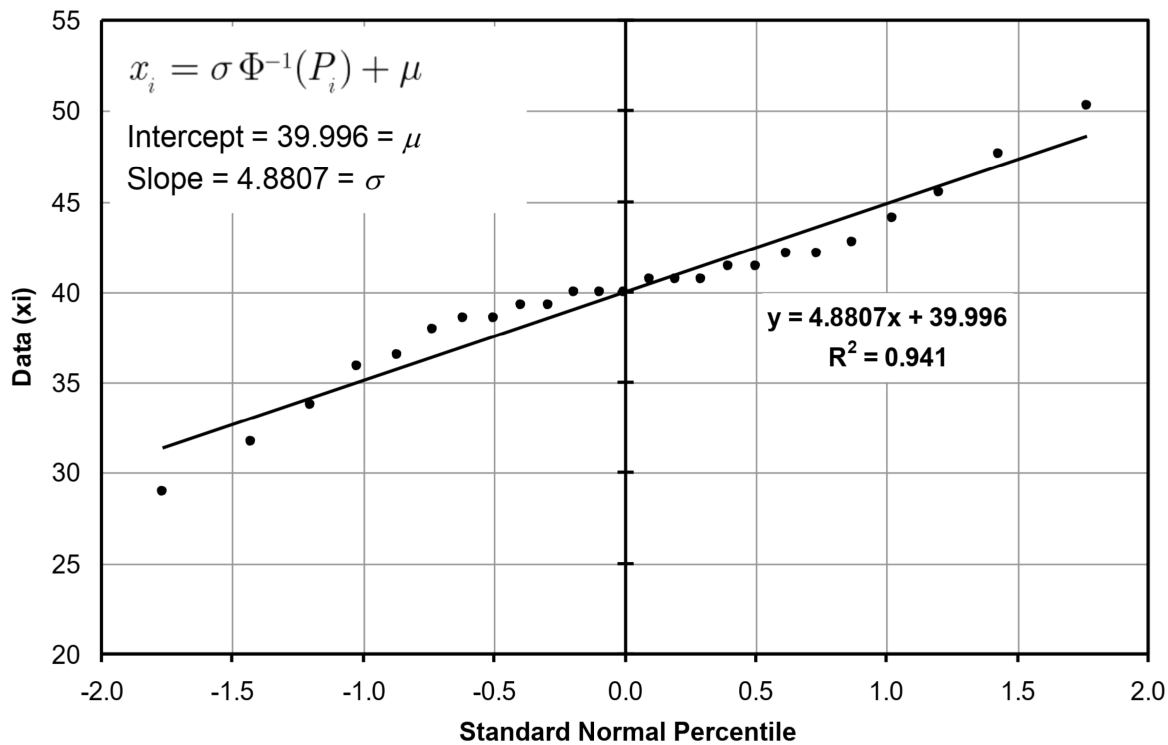
– *Step 3:* Calculate the corresponding percentile value.

– *Step 4:* Plot the data.

– *Step 5:* Fit a straight line through the data.

Test No.	Crushing Strength (MPa)	(Step 1) Rank (i)	(Step 2) P_i	(Step 3) $\Phi^{-1}(P_i)$
3	29.0	1	0.0385	-1.7688
4	31.7	2	0.0769	-1.4261
10	33.8	3	0.1154	-1.1984
17	35.9	4	0.1538	-1.0201
24	36.5	5	0.1923	-0.8694
2	37.9	6	0.2308	-0.7363
16	38.6	7	0.2692	-0.6151
23	38.6	8	0.3077	-0.5024
5	39.3	9	0.3462	-0.3957
11	39.3	10	0.3846	-0.2934
...	...	...	...	...

Example: Normal PPP (cont'd)



Lognormal Probability Paper Plots

- Equate the rank probability of the data with the lognormal CDF:

$$P_i = F(x_i) = \Phi\left(\frac{\ln x_i - \theta}{\omega}\right)$$

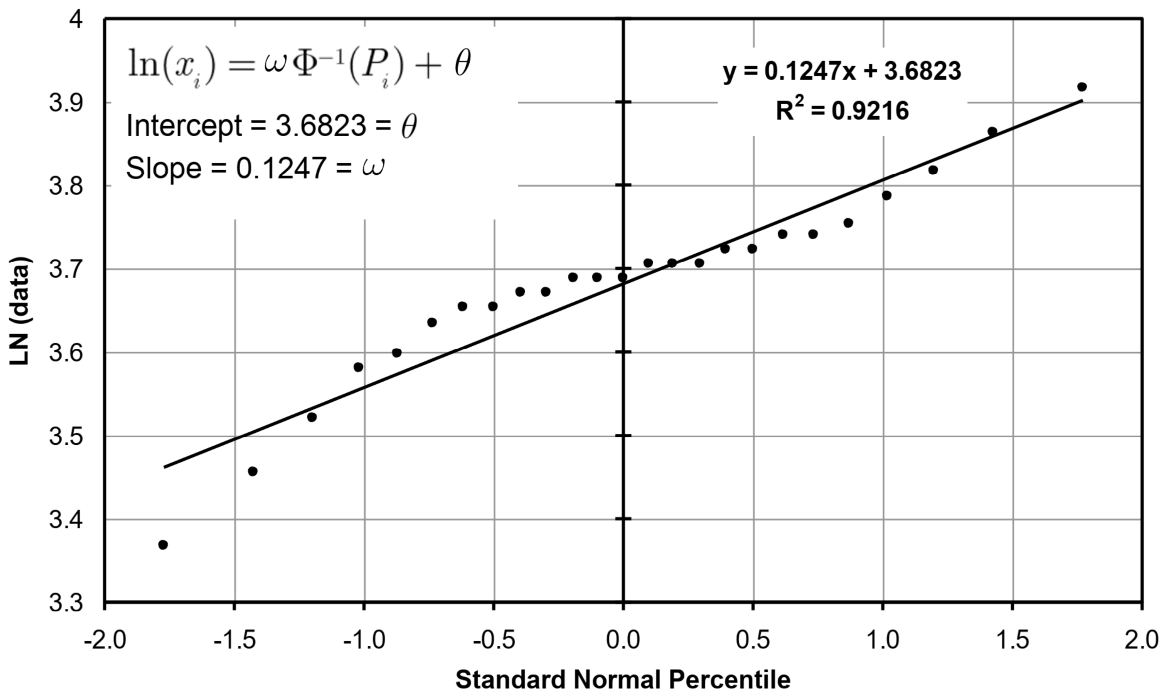
$$\implies \Phi^{-1}(P_i) = \frac{\ln x_i - \theta}{\omega} \quad (\text{replace: } P_i = \tfrac{i}{n+1})$$

$$\implies \ln x_i = \omega\Phi^{-1}\left(\frac{i}{n+1}\right) + \theta \quad (\text{form: } y = mx + c)$$

  where $\Phi(\cdot)$ is the standard normal CDF; $\Phi^{-1}(\cdot)$ is the standard normal inverse CDF.

- The lognormal PPP is constructed by plotting the logarithm of data $\ln(x_i)$ against the corresponding percentile value given by $\Phi^{-1}(P_i)$.
- Distribution parameters can be estimated from the above straight line equation as: θ = intercept and ω = slope.

Example: Lognormal PPP



Weibull Probability Paper Plots

– Equate the rank probability of the data with the Weibull CDF:

$$P_i = F(x_i) = 1 - e^{-\left(\frac{x_i}{\theta}\right)^\beta} \implies \ln(1 - P_i) = -\left(\frac{x_i}{\theta}\right)^\beta$$

$$\implies \frac{1}{\beta}\ln(-\ln(1 - P_i)) = \ln x_i - \ln\theta$$

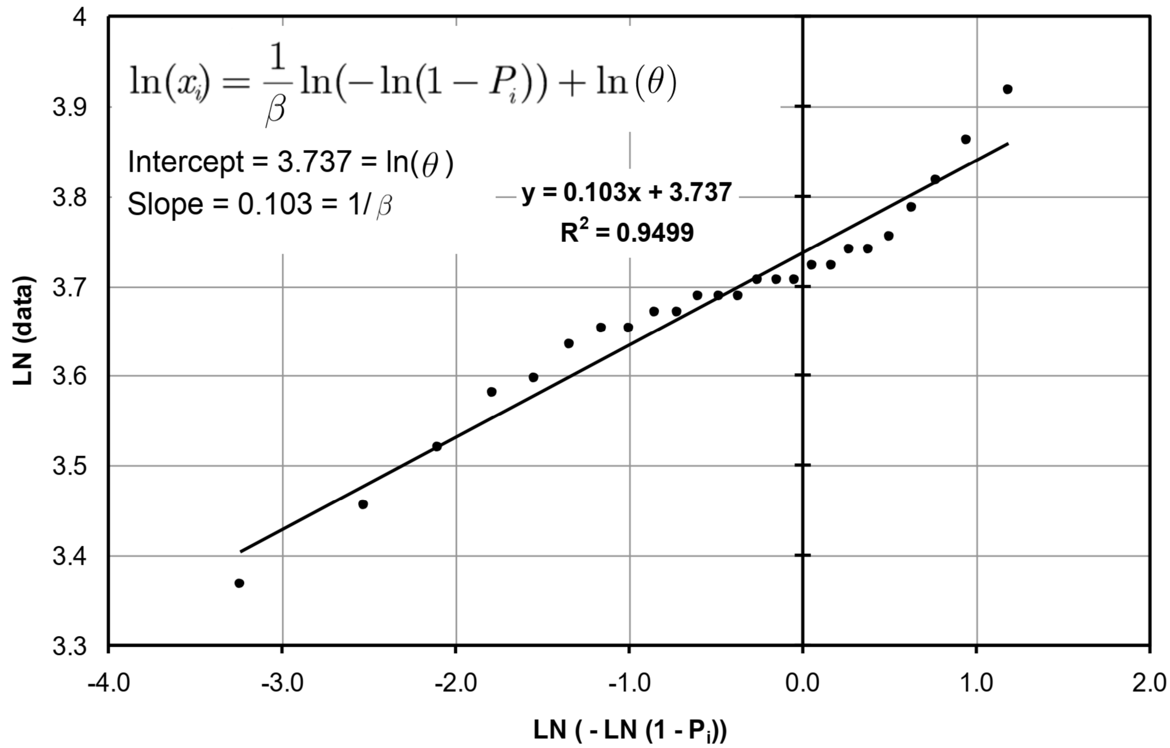
$$\implies \ln x_i = \frac{1}{\beta}\ln(-\ln(1 - P_i)) + \ln\theta \quad (\text{replace: } P_i = \tfrac{i}{n+1})$$

$$\implies \ln x_i = \frac{1}{\beta}\ln\left(-\ln\left(1 - \frac{i}{n+1}\right)\right) + \ln\theta \quad (\text{form: } y = mx + c)$$

– The Weibull PPP is constructed by plotting the logarithm of data $\ln(x_i)$ against the corresponding value given by $\ln(-\ln(1 - P_i))$.

– The distribution parameters can be estimated as: $\beta = \frac{1}{\text{slope}}$ and $\theta = e^{\text{intercept}}$

Example: Weibull PPP (cont'd)



Exponential Probability Paper Plots

– Equate the rank probability of the data with the exponential CDF:

$$
\begin{aligned}
& P_i = F(x_i) = 1 - e^{-\lambda x_i} \\
\Longrightarrow\ & \ln(1-P_i) = -\lambda x_i \\
\Longrightarrow\ & x_i = -\frac{1}{\lambda}\ln(1-P_i) \quad (\text{replace: } P_i = \tfrac{i}{n+1}) \\
\Longrightarrow\ & x_i = -\frac{1}{\lambda}\ln\left(1 - \frac{i}{n+1}\right) \quad (\text{form: } y = mx + 0)
\end{aligned}
$$

– The exponential PPP is constructed by plotting the data x_i against the corresponding value given by $\{-\ln(1-P_i)\}$.

– The distribution parameter can be estimated as: $\lambda = \frac{1}{\text{slope}}$

Maximum Likelihood Estimation (MLE)

MLE is a statistical method used to estimate the parameters of a statistical model. The goal of MLE is to find the values of the model parameters that *maximize the likelihood function*, which measures how well the model explains the observed data.

- Assume that the sample data $x_1, x_2, \cdots, x_n$ follow a probability distribution with PDF $f(x|\theta_1, \theta_2 \cdots, \theta_k)$, where $\theta_{1:k}$ are the distribution parameters.

- The likelihood of obtaining these n independent data points from the chosen distribution can be mathematically expressed as

$$\begin{aligned} L(\theta_{1:k}|x_{1:n}) &= f(x_1|\theta_{1:k}) \times f(x_2|\theta_{1:k}) \times \cdots \times f(x_n|\theta_{1:k}) \\ &= \prod_{i=1}^{n} f(x_i|\theta_{1:k}) \end{aligned}$$

[7] Distribution Selection • § Maximum Likelihood Estimation
© 2025 Indranil Hazra

Maximum Likelihood Estimation (MLE) (cont'd)

- The distribution parameters can be estimated by **maximizing** the **likelihood** function $L(\theta_{1:k}|x_{1:n})$ or by maximizing the **logarithm of the likelihood** function $\ln L(\theta_{1:k}|x_{1:n})$.

- For estimating k number of parameters, a system of k equations can be obtained by maximizing the log-likelihood function as

$$\frac{\partial(\ln L(\theta_{1:k}|x_{1:n}))}{\partial\theta_1} = 0 \quad \text{Equation (1)}$$

$$\frac{\partial(\ln L(\theta_{1:k}|x_{1:n}))}{\partial\theta_2} = 0 \quad \text{Equation (2)}$$

$$\vdots$$

$$\frac{\partial(\ln L(\theta_{1:k}|x_{1:n}))}{\partial\theta_k} = 0 \quad \text{Equation (k)}$$

MLE for Exponential Distribution

– Exponential PDF:

$$f(x) = \lambda e^{-\lambda x}, \quad x \geq 0$$

– Likelihood function:

$$L(\lambda | x_{1:n}) = \prod_{i=1}^{n} f(x_i | \lambda) = \lambda^n \prod_{i=1}^{n} e^{-\lambda x_i}$$

– Log-likelihood function:

$$\ln L(\lambda | x_{1:n}) = n \ln \lambda - \lambda \sum_{i=1}^{n} x_i$$

– Parameter estimation by maximizing the log-likelihood:

$$\frac{d \ln L(\lambda | x_{1:n})}{d\lambda} = 0 \implies \frac{n}{\lambda} - \sum_{i=1}^{n} x_i = 0 \implies \lambda = \frac{n}{\sum_{i=1}^{n} x_i}$$

MLE for Normal Distribution

– Normal PDF:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right], \quad -\infty < x < \infty$$

– Likelihood function:

$$L(\mu, \sigma | x_{1:n}) = \prod_{i=1}^{n} f(x_i | \mu, \sigma) = \frac{1}{\sigma^n (2\pi)^{n/2}} \prod_{i=1}^{n} e^{-\frac{1}{2}\left(\frac{x_i - \mu}{\sigma}\right)^2}$$

– Log-likelihood function:

$$\ln L(\mu, \sigma | x_{1:n}) = -n \ln \sigma - \frac{n}{2} \ln 2\pi - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2$$

MLE for Normal Distribution (cont'd)

- Parameter estimation by maximizing the log-likelihood for μ:

$$\frac{\partial \ln L(\mu, \sigma|x_{1:n})}{\partial \mu} = 0$$

$$\implies + \frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i - \mu) = 0$$

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

- Parameter estimation by maximizing the log-likelihood for σ:

$$\frac{\partial \ln L(\mu, \sigma|x_{1:n})}{\partial \sigma} = 0$$

$$\implies - \frac{n}{\sigma} + \frac{1}{\sigma^3} \sum_{i=1}^{n} (x_i - \mu)^2 = 0$$

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

MLE for Lognormal Distribution

- Lognormal PDF:

$$f(x) = \frac{1}{x\omega\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\ln x - \theta}{\omega}\right)^2 \right], \quad 0 < x < \infty$$

- Likelihood function:

$$L(\theta, \omega|x_{1:n}) = \prod_{i=1}^{n} f(x_i|\theta, \omega) = \frac{1}{\omega^n (2\pi)^{n/2}} \prod_{i=1}^{n} \frac{e^{-\frac{1}{2}\left(\frac{\ln x_i - \theta}{\omega}\right)^2}}{x_i^n}$$

- Log-likelihood function:

$$\ln L(\theta, \omega|x_{1:n}) = -n \ln \omega - \frac{n}{2} \ln 2\pi - \frac{1}{2\omega^2} \sum_{i=1}^{n} (\ln x_i - \theta)^2 - n \sum_{i=1}^{n} \ln x_i$$

MLE for Lognormal Distribution (cont'd)

– Parameter estimation by maximizing the log-likelihood for θ:

$$\frac{\partial \ln L(\theta, \omega | x_{1:n})}{\partial \theta} = 0$$
$$\implies + \frac{1}{\omega^2} \sum_{i=1}^{n} (\ln x_i - \theta) = 0$$
$$\implies \theta = \frac{\sum_{i=1}^{n} \ln x_i}{n}$$

– Parameter estimation by maximizing the log-likelihood for ω:

$$\frac{\partial \ln L(\theta, \omega | x_{1:n})}{\partial \omega} = 0$$
$$\implies -\frac{n}{\omega} + \frac{1}{\omega^3} \sum_{i=1}^{n} (\ln x_i - \theta)^2 = 0$$
$$\implies \omega^2 = \frac{\sum_{i=1}^{n} (\ln x_i - \theta)^2}{n}$$

MLE for Gamma Distribution

– Gamma PDF:

$$f(x) = \frac{\lambda (\lambda x)^{r-1} e^{-\lambda x}}{\Gamma(r)}, \quad x \geq 0$$

– Likelihood function:

$$L(r, \lambda | x_{1:n}) = \prod_{i=1}^{n} f(x_i | r, \lambda) = \frac{\lambda^{nr}}{\Gamma(r)^n} \prod_{i=1}^{n} x_i^{r-1} e^{-\lambda x_i}$$

– Log-likelihood function:

$$\ln L(r, \lambda | x_{1:n}) = nr \ln \lambda - n \ln \Gamma(r) + (r-1) \sum_{i=1}^{n} \ln x_i - \lambda \sum_{i=1}^{n} x_i$$

MLE for Gamma Distribution (cont'd)

– Parameter estimation by maximizing the log-likelihood for λ:

$$\frac{\partial \ln L(r, \lambda | x_{1:n})}{\partial \lambda} = 0$$

$$\implies \frac{nr}{\lambda} - \sum_{i=1}^{n} x_i = 0 \implies \lambda = \frac{nr}{\sum_{i=1}^{n} x_i}$$

– Parameter estimation by maximizing the log-likelihood for r:

$$\frac{\partial \ln L(r, \lambda | x_{1:n})}{\partial r} = 0$$

$$\implies n \ln \lambda - n \frac{d \ln \Gamma(r)}{dr} + \sum_{i=1}^{n} \ln x_i = 0 \quad \text{(replace } \lambda\text{)}$$

$$\implies \ln r - \psi(r) = \ln \frac{\sum_{i=1}^{n} x_i}{n} - \frac{\sum_{i=1}^{n} \ln x_i}{n}$$

$$\left(\text{digamma function: } \psi(r) = \tfrac{d \ln \Gamma(r)}{dr}\right)$$

– *Can you solve it analytically? No!* Use a numerical solver.

[7] Distribution Selection • § Maximum Likelihood Estimation

MLE for Weibull Distribution

– Weibull PDF:

$$f(x) = \frac{\beta}{\theta}\left(\frac{x}{\theta}\right)^{\beta-1} \exp\left[-\left(\frac{x}{\theta}\right)^{\beta}\right], \quad x \geq 0$$

– Likelihood function:

$$L(\beta, \theta | x_{1:n}) = \prod_{i=1}^{n} f(x_i | \beta, \theta) = \frac{\beta^n}{\theta^{n\beta}} \prod_{i=1}^{n} x_i^{\beta-1} e^{-(\frac{x_i}{\theta})^{\beta}}$$

– Log-likelihood function:

$$\ln L(\beta, \theta | x_{1:n}) = n \ln \beta - n\beta \ln \theta + (\beta - 1) \sum_{i=1}^{n} \ln x_i - \frac{1}{\theta^{\beta}} \sum_{i=1}^{n} {x_i}^{\beta}$$

MLE for Weibull Distribution (cont'd)

– Parameter estimation by maximizing the log-likelihood for θ:

$$\frac{\partial \ln L(\beta,\theta|x_{1:n})}{\partial\theta} = 0 \implies -\frac{n\beta}{\theta} + \frac{\beta}{\theta^{\beta+1}}\sum_{i=1}^{n} {x_i}^{\beta} = 0 \implies \theta = \left(\frac{\sum_{i=1}^{n} {x_i}^{\beta}}{n}\right)^{\frac{1}{\beta}}$$

– Parameter estimation by maximizing the log-likelihood for β:

$$\frac{\partial \ln L(\beta,\theta|x_{1:n})}{\partial\beta} = 0$$

$$\implies \frac{n}{\beta} - n\ln\theta + \sum_{i=1}^{n}\ln x_i - \frac{1}{\theta^{\beta}}\sum_{i=1}^{n} {x_i}^{\beta}\ln x_i + \frac{\ln\theta}{\theta^{\beta}}\sum_{i=1}^{n} {x_i}^{\beta} = 0$$

$$\implies \frac{n}{\beta} - n\ln\theta + \sum_{i=1}^{n}\ln x_i - \frac{n\sum_{i=1}^{n} {x_i}^{\beta}\ln x_i}{\sum_{i=1}^{n} {x_i}^{\beta}} + n\ln\theta = 0 \quad \text{(replace } \theta\text{)}$$

$$\implies \frac{1}{\beta} + \frac{1}{n}\sum_{i=1}^{n}\ln x_i - \frac{\sum_{i=1}^{n} {x_i}^{\beta}\ln x_i}{\sum_{i=1}^{n} {x_i}^{\beta}} = 0 \quad \text{(Use a numerical solver!)}$$

MLE Example

– Data on aluminium contamination (ppm) in plastic:

Aluminum Contamination (ppm)

30	30	60	63	70	79	87
90	101	102	115	118	119	119
120	125	140	145	172	182	
183	191	222	244	291	511	

From "The Lognormal Distribution for Modeling Quality Data When the Mean Is Near Zero," *Journal of Quality Technology*, 1990, pp. 105–110.

– Sample mean, $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = 142.6538$ ppm, where $n = 26$.

– Sample variance, $s^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1} = 98.2043^2$.

MLE Example (cont'd)

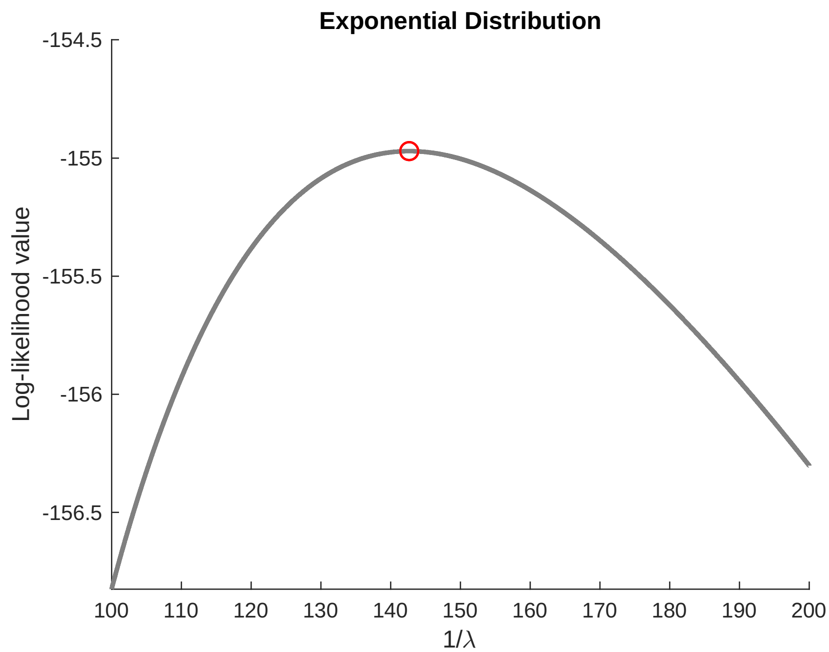


Figure: The **log-likelihood profile** for the **exponential** distribution fitted to the aluminium contamination data.

[7] Distribution Selection • § Maximum Likelihood Estimation

© 2025 Indranil Hazra

MLE Example (cont'd)

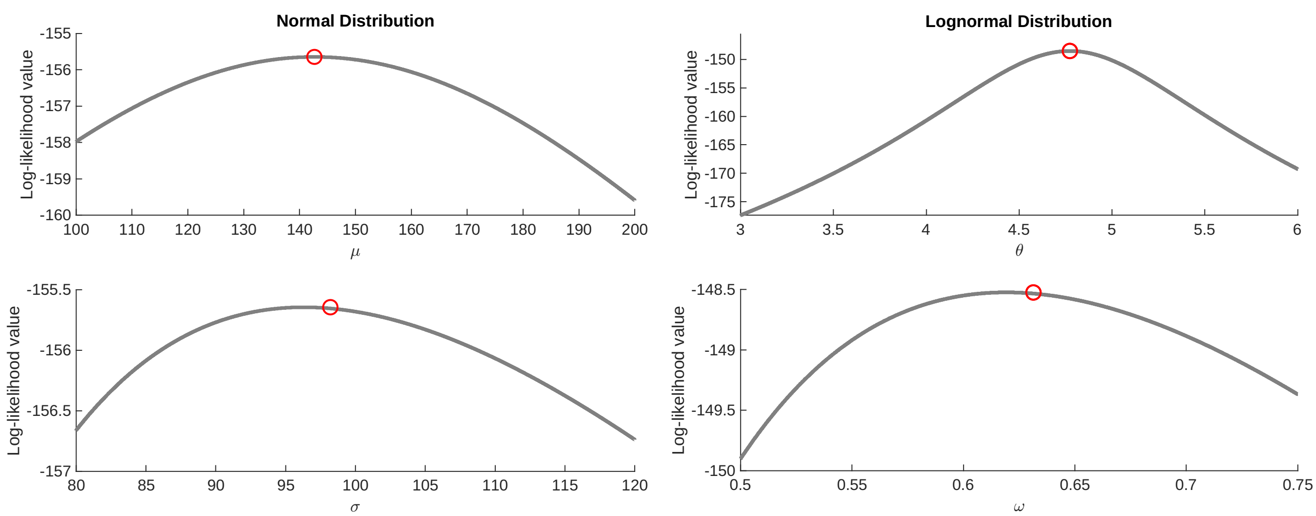


Figure: The **log-likelihood profiles** for **normal** (*left*) and **lognormal** (*right*) distributions fitted to the aluminium contamination data.

MLE Example (cont'd)

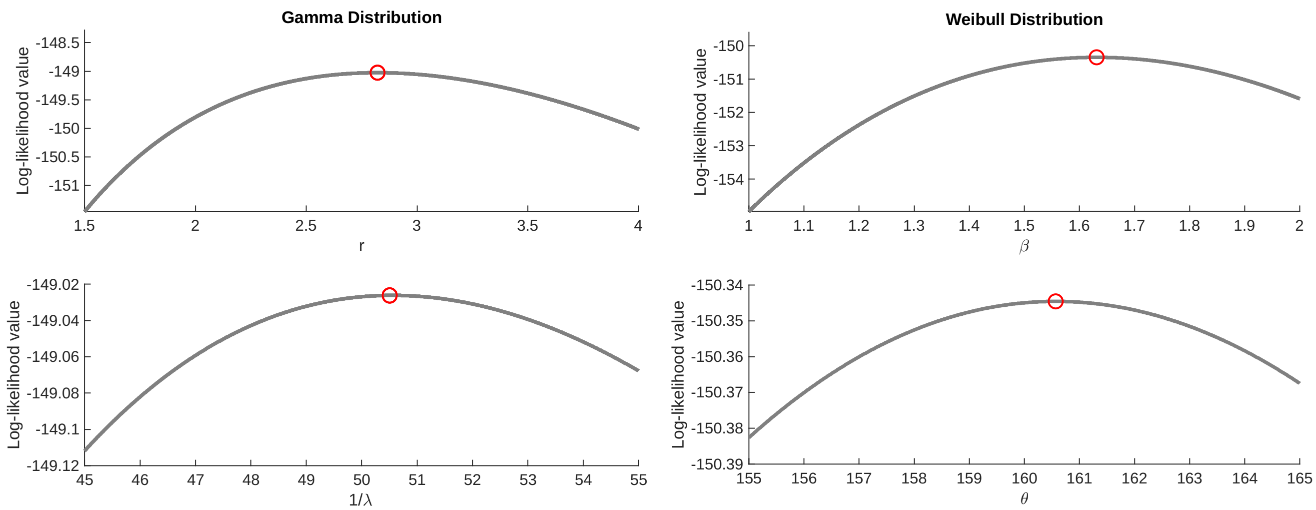


Figure: The **log-likelihood profiles** for **gamma** (*left*) and **Weibull** (*right*) distributions fitted to the aluminium contamination data.

MLE Example (cont'd)

Table: Maximum likelihood (ML) estimates of distribution parameters.

Prob. Dist.	Parameter	ML estimate	Log-likelihood	MoM estimate
Exponential	λ	$\hat{\lambda}_{mle} = 0.0070$	-154.9709	$\hat{\lambda}_{mom} = 0.0070$
Normal	μ, σ	$\hat{\mu}_{mle} = 142.654$ $\hat{\sigma}_{mle} = 98.204$	-155.6458	$\hat{\mu}_{mom} = 142.654$ $\hat{\sigma}_{mom} = 98.204$
Lognormal	θ, ω	$\hat{\theta}_{mle} = 4.7729$ $\hat{\omega}_{mle} = 0.6314$	-148.5235	$\hat{\theta}_{mom} = 4.7665$ $\hat{\omega}_{mom} = 0.6228$
Gamma	r, λ	$\hat{r}_{mle} = 2.8215$ $\hat{\lambda}_{mle} = 0.0198$	-149.0262	$\hat{r}_{mom} = 2.1101$ $\hat{\lambda}_{mom} = 0.0148$
Weibull	β, θ	$\hat{\beta}_{mle} = 1.6312$ $\hat{\theta}_{mle} = 160.57$	-150.3446	$\hat{\beta}_{mom} = 1.478$ $\hat{\theta}_{mom} = 157.727$

MLE Example (cont'd)

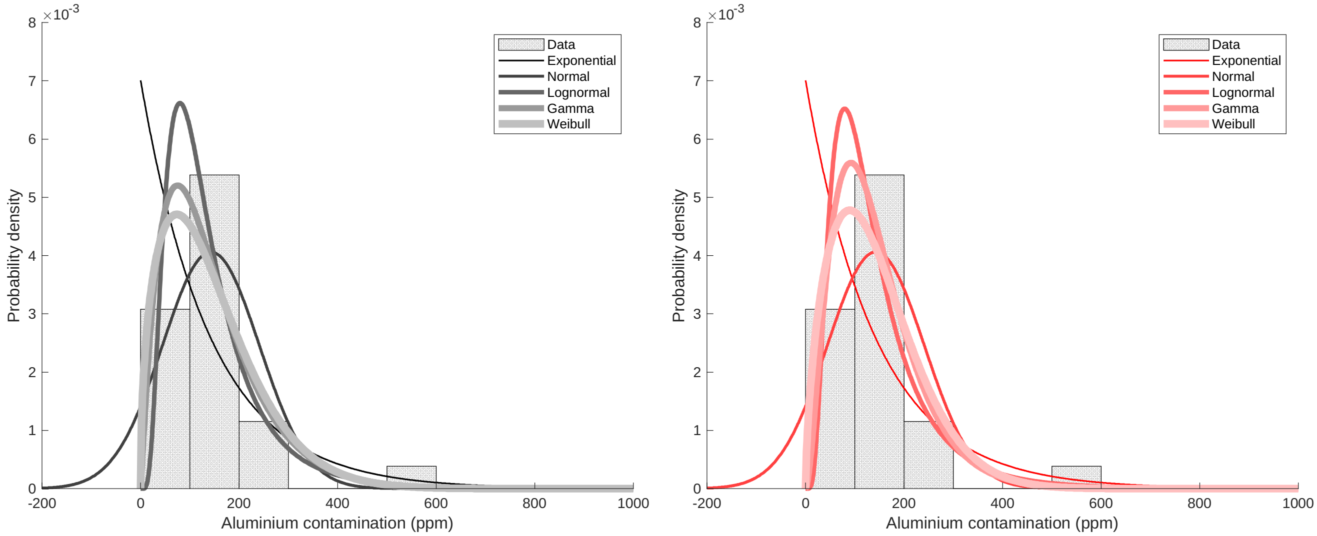


Figure: Various distributions fitted to the aluminium contamination data using **method of moments** (*left*) and **method of maximum likelihood** (*right*).