$$X_i \overset{iid}{\sim} \text{Gamma}(\alpha, \beta)$$

$$\underline{Y_i} = \log(X_i)$$

$$\left\{\begin{aligned} m_1 &= \frac{1}{n}\sum_i \log(X_i) \\ m_2 &= \frac{1}{n}\sum_i (\log X_i)^2 \end{aligned}\right.$$

$$M_X(t) = E e^{Xt}$$

$$C_X(t) = \log E e^{Xt} \rightarrow \text{cumulant generating function}$$

$$EX = (C_X(t))'\Big|_{t=0} \qquad V(X) = (C_X(t))''\Big|_{t=0}$$

$$C_Y(t) \qquad M_Y(t) = \int \underline{f_Y(y)}\, e^{yt}\, dy$$

$$\underline{M_Y(t)} = E e^{Yt} = E e^{t \cdot \ln X} = E X^t = \int_0^\infty x^t \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}} dx$$

$$= \frac{1}{\beta^\alpha \Gamma(\alpha)} \int_0^\infty \underline{x^{t+\alpha-1}} e^{-\frac{x}{\beta}} dx = \frac{\beta^{\alpha+t}\,\Gamma(\alpha+t)}{\beta^\alpha \Gamma(\alpha)} \int_0^\infty \frac{x^{t+\alpha-1} e^{-\frac{x}{\beta}}}{\beta^{t+\alpha}\Gamma(\alpha+t)} dx$$

$$= \beta^t \frac{\Gamma(\alpha+t)}{\Gamma(\alpha)}$$

$$C_Y(t) = t\,\underline{\ln\beta} + \underline{\ln\Gamma(\alpha+t)} - \ln\Gamma(\alpha)$$

$$EY = \frac{d\,C_Y(t)}{dt}\Big|_{t=0} = (\underline{\ln\beta + \psi(\alpha+t)})\Big|_{t=0} = \ln\beta + \psi(\alpha)$$

$$m_2 - m_1^2 \quad V(Y) = \frac{d^2 c_Y(t)}{dt^2}\Big|_{t=0} = \psi_1(\alpha+t)\Big|_{t=0} = \psi_1(\alpha)$$

Theorem 5.3.1 Assume that $X_1, X_2 \cdots X_n \overset{iid}{\sim} f(x)$ with mean of $\mu$ and variance of $\sigma^2 < \infty$

$$m_1 = \bar{X} = \mu$$

$$\Rightarrow \hat{\mu} = \bar{X}$$

① $\hat{\mu}$ is unbiased.

② $\hat{\mu}$ is consistent

In general. MOM is always consistent.

MOM is not unbiased.

Trafic Light:

$$X_1 \cdots X_n \qquad \underline{14 + 30 + 36 + \cdots + 1 + 1 + 0}$$

$$X_i \quad 0, 1, 2, \cdots 12, 13+$$

$$\underbrace{0 \cdots 0}_{14} \quad \underbrace{1 \cdots 1}_{30} \quad \underbrace{2 \sim 2}_{36} \quad \cdots$$

$$X_i \sim \text{Poisson}(\lambda)$$

$$m_1 = \bar{X} \qquad EX = \lambda \qquad m_2 = \frac{1}{n}\Sigma X_i^2$$

$$EX^2 = \lambda + \lambda^2$$

① $\hat{\lambda} = \bar{X}$

$$\lambda + \lambda^2 = m_2$$

$$\hat{\lambda} \overset{\text{||}}{=} \lambda_+$$

MOM: ① estimators are not unique.

② Sometimes the MOM estimator doesn't make sense

## MLE (Maximum Likelihood Estimate)

Likelihood function: Joint density of the data, viewed as a function of the parameter

$$X \sim Exp(\theta)$$

$$f(x;\theta) = \frac{1}{\theta} e^{-\frac{x}{\theta}}$$

$$L(\theta; x) = \frac{1}{\theta} e^{-\frac{x}{\theta}}$$

MLE: want to find $\hat{\theta}$ such that $L(\theta; x)$ achieves the maximum

| $\theta$ | Y: -1 | 0 | 1 |
|---|---|---|---|
| 1 | 0.2 | 0.3 | 0.5 |
| 2 | 0.7 | 0.2 | 0.1 |
| 3 | 0.2 | 0.6 | 0.2 |

when $Y=0$ MLE $\hat{\theta} = 3$

In general: $X_1 \dots X_n \overset{iid}{\sim} f(x|\theta)$ Joint density $\prod_{i=1}^{n} f(y_i|\theta)$

$$\underline{L_n}(\theta) = \prod_{i=1}^{n} f(y_i|\theta), \quad \text{log-likelihood } l(\theta) = \log L_n(\theta) = \sum_{i=1}^{n} \log f(y_i|\theta)$$

$$\hat{\theta} = \operatorname{argmax}\{\theta : L_n(\theta)\}$$

$i=1$

$$\hat{\theta} = \underset{\theta}{\operatorname{argmax}} \{ \theta, \ L_n(\theta) \}$$

GDP Example: $Y_1, Y_2, \cdots, Y_n \overset{iid}{\sim} N(\mu, \sigma^2)$ $\quad \theta = (\mu, \sigma^2)$

$$f(y_i \mid \theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ -\frac{(y_i-\mu)^2}{2\sigma^2} \right\}$$

$$L(\theta; \tilde{y}) = \prod f(y_i \mid \theta) = \left( \frac{1}{\sqrt{2\pi\sigma^2}} \right)^n \exp\left\{ -\frac{\Sigma(x_i-\mu)^2}{2\sigma^2} \right\}$$

$$\ell(\theta; \tilde{y}) = \underbrace{-n \log\left(\sqrt{2\pi\sigma^2}\right)}_{-\frac{n}{2}\log(2\pi\sigma^2)} - \frac{\Sigma(x_i-\mu)^2}{2\sigma^2} \qquad \left(\frac{1}{\sigma^2}\right) = (\sigma^2)^{-1}$$

$$\frac{\partial \ell}{\partial \mu} = -\frac{1}{2\sigma^2} \sum 2(x_i-\mu)(-1) = \frac{1}{\sigma^2}\Sigma(x_i-\mu) = 0$$

$\Sigma x_i - \Sigma \mu$

$\Sigma x_i - n\mu$

$$\frac{\partial \ell}{\partial \sigma^2} = -\frac{n}{2}\frac{\cancel{2\pi}}{\cancel{2\pi}\sigma^2} - \frac{\Sigma(x_i-\mu)^2}{2}(-1)\cdot(\sigma^2)^{-2}$$

$$= -\frac{n\sigma^2}{2\cancel{\sigma^4}} + \frac{\Sigma(x_i-\mu)^2}{2\cancel{(\sigma^2)^2}} = 0$$

$$\begin{cases} \hat{\mu} = \bar{x} \\ \hat{\sigma}^2 = \dfrac{\Sigma(x_i-\mu)^2}{n} = \dfrac{\Sigma(x_i-\bar{x})^2}{n} \end{cases}$$

For normal case, MLE for $\mu$ and $\sigma^2$ is the same as that based on the MoM.

What is the MLE for the Log-Normal?

~~lo~~ $X \sim$ Lognormal $\qquad \log X$

MOM $\quad \hat{\mu}_1 \quad \hat{\sigma}_1^2 \qquad \hat{\mu}_2 \quad \hat{\sigma}_2^2$

$\hat{\mu}$

ML =

$\hat{\sigma}^2$

$\hat{\mu}$

MLE $\hat{\sigma}^2$

MLE is consistent

Example 5.4.4. Traffic light.

$X_i \overset{iid}{\sim} \text{Poisson}(\lambda)$

$$f(x_i | \lambda) = \frac{\lambda^{x_i}}{x_i!} e^{-\lambda}$$

$$L(\lambda; \tilde{x}) = \prod \frac{\lambda^{x_i}}{x_i!} e^{-\lambda} = e^{-n\lambda} \frac{\lambda^{\sum x_i}}{\prod x_i!}$$

$$l(\lambda; \tilde{x}) = -n\lambda + \sum x_i \ln \lambda - \ln (\prod x_i!)$$

$$\frac{dl}{d\lambda} = -n + \frac{\sum x_i}{\lambda} = 0 \Rightarrow \hat{\lambda} = \frac{\sum x_i}{n} = \bar{X}$$

Example 5.4.3 Let $X_i$ be the number of damaged O-rings

$X_i \overset{iid}{\sim} \text{Bin}(2, p)$

$$f(x_i | p) = \binom{2}{x_i} p^{x_i} (1-p)^{2-x_i}$$

$$L(p; \tilde{x}) = \prod \binom{2}{x_i} p^{\sum x_i} (1-p)^{2n - \sum x_i}$$

$$l(p; \tilde{x}) = \log \left( \prod \binom{2}{x_i} \right) + \sum x_i \ln p + (2n - \sum x_i) \ln (1-p)$$

$$\frac{dl}{dp} = \frac{\sum x_i}{p} - \frac{2n - \sum x_i}{1-p} = 0$$

$$\frac{\sum x_i}{p} = \frac{2n - \sum x_i}{1-p} \Rightarrow \sum x_i - \cancel{p \sum x_i} = 2np - \cancel{p \sum x_i}$$

$$\hat{p} = \frac{\sum x_i}{2n}$$

$$x_i \sim \text{Bin}(2, \; p_i)$$

logistic Regression

$p_i$ relates to temperature $t_i$

$$\log \frac{p_i}{1-p_i} = \beta_0 + \beta_1 t_i$$

Two Parameters

MLE

$$\frac{p_i}{1-p_i} = \exp(\beta_0 + \beta_1 t_i)$$

$$p_i = \frac{\exp(\beta_0 + \beta_1 t_i)}{1+\exp(\beta_0 + \beta_1 t_i)}$$

$$f(x_i \mid p_i) = \binom{2}{x_i} p_i^{x_i} (1-p_i)^{2-x_i}$$

$$= \binom{2}{x_i} \frac{(\exp(\beta_0 + \beta_1 t_i))^{x_i}}{(1+\exp(\beta_0+\beta_1 t_i))^{x_i}} \frac{1}{(1+\exp(\beta_0+\beta_1 t_i))^{2-x_i}}$$

$$= \binom{2}{x_i} \frac{(\exp(\beta_0 + \beta_1 t_i))^{x_i}}{(1+\exp(\beta_0+\beta_1 t_i))^{2}}$$

$$L(\beta_0, \beta_1) = \prod \binom{2}{x_i} \prod_{i=1}^{n} \frac{(\exp(\beta_0 + \beta_1 t_i))^{x_i}}{(1+\exp(\beta_0+\beta_1 t_i))^{2}}$$

$$\ell(\beta_0, \beta_1) = \ln \prod \binom{2}{x_i} + \sum_{i=1}^{n} \left( x_i(\beta_0 + \beta_1 t_i) - 2\ln(1+\exp(\beta_0+\beta_1 t_i)) \right)$$

$$\begin{cases} \dfrac{\partial \ell}{\partial \beta_0} = 0 \\ \dfrac{\partial \ell}{\partial \beta_1} = 0 \end{cases} \Rightarrow \text{No explicit formula!}$$

$$\frac{\partial l}{\partial \beta_1} = 0 \quad \Rightarrow \text{No explicit formula !!}$$

Newton Raphson Algorithm:

GDP $X_i \sim \text{Gamma}(\alpha, \beta)$

$$f(x_i \mid \alpha, \beta) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}}$$

$$L(\alpha, \beta; \tilde{x}) = \left(\frac{1}{\beta^\alpha \Gamma(\alpha)}\right)^n \left(\prod x_i\right)^{\alpha-1} e^{-\frac{\sum x_i}{\beta}}$$

$$l(\alpha, \beta; \tilde{x}) = -n\left(\alpha \ln\beta + \ln\Gamma(\alpha)\right) + (\alpha-1)\sum \ln x_i - \frac{\sum x_i}{\beta}$$

$$\begin{cases} \dfrac{\partial l}{\partial \alpha} = -n\left(\ln\beta + \psi(\alpha)\right) + \sum \ln x_i = 0 \\ \dfrac{\partial l}{\partial \beta} = -n\left(\dfrac{\alpha}{\beta}\right) + \dfrac{\sum x_i}{\beta^2} = 0 \end{cases}$$

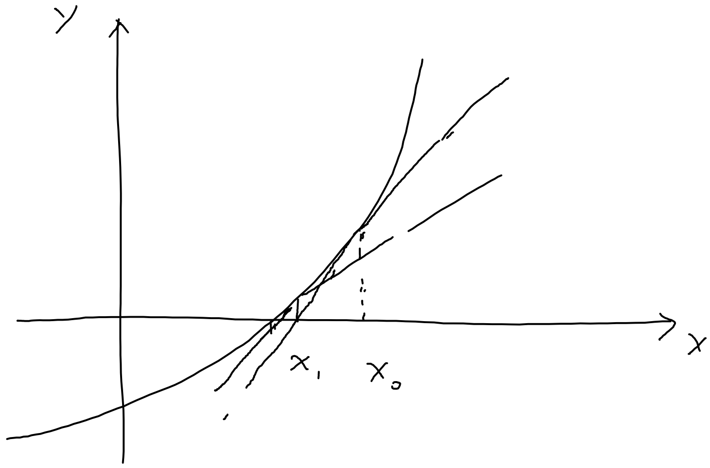
$$\begin{cases} \ln\beta + \psi(\alpha) - \frac{1}{n}\sum \ln x_i = 0 \\ \alpha\beta - \frac{1}{n}\sum x_i = 0 \end{cases}$$

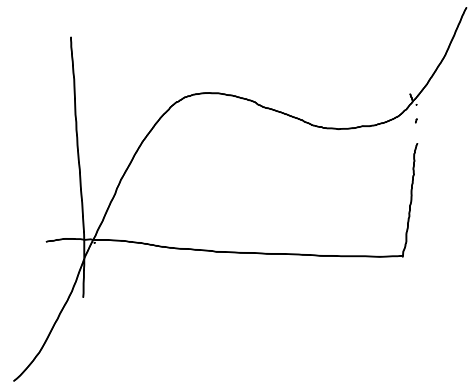
Newton Raphson algorithm

$$f(x) = 0$$

1. Choose an initial value $x_0$
2. Assume $x_k$, $x_{k+1} = x_k - \dfrac{f(x_k)}{f'(x_k)}$
3. $\Delta = |x_{k+1} - x_k|$
4. If $\Delta > \delta$ Repeat 2 and 3; Otherwise, stop

$$\hat{x} = x_{k+1}$$



HW. $\psi_1(\alpha) = m_2 - m_1^2$

$$f(\alpha) = \psi_1(\alpha) - (m_2 - m_1^2)$$

Choose an initial value

② $\alpha_k, \alpha_k, \alpha_{k+1} = \alpha_k - \frac{f(\alpha_k)}{f'(\alpha_k)} = \alpha_k -$

$$\alpha_k \quad \alpha_{k+1} = \alpha_k - \frac{f(\alpha_k)}{f'(\alpha_k)} = \alpha_k - \frac{\psi_1(\alpha_k) - (m_2 - m_1^2)}{\psi_1'(\alpha_k)}$$

$$\begin{cases} f_1(x_1, x_2 \cdots x_n) = 0 \\ f_2(x_1, x_2 \cdots x_n) = 0 \\ \vdots \\ f_n(x_1, x_2, \cdots x_n) = 0 \end{cases}$$

$$\tilde{x}^i \quad \tilde{x}^{i+1} = \tilde{x}^i - \left(J(\tilde{x}^i)\right)^{-1} \tilde{f}(\tilde{x}^i)$$

$$J(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \end{pmatrix}$$

$$J(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_n} \\ \vdots & & \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_n} \end{pmatrix}$$

(Gamma MLE)

$$\begin{cases} f_1(\alpha, \beta) = \frac{1}{n} \sum \ln x_i - \underline{\underline{\psi}}(\alpha) - \ln \beta \\ f_2(\alpha, \beta) = \frac{1}{n} \sum x_i - \alpha \underline{\beta} \end{cases}$$

$$\begin{pmatrix} \alpha^{i+1} \\ \beta^{i+1} \end{pmatrix} = \begin{pmatrix} \alpha^i \\ \beta^i \end{pmatrix} - \begin{pmatrix} -\psi_1(\alpha^i) & -\frac{1}{\beta^i} \\ -\beta^i & -\alpha^i \end{pmatrix}^{-1} \begin{pmatrix} f_1(\alpha^i, \beta^i) \\ f_2(\alpha^i, \beta^i) \end{pmatrix}$$

$$J = \begin{pmatrix} -\psi_1(\alpha) & -\frac{1}{\beta} \\ -\beta & -\alpha \end{pmatrix}$$

$$\Delta = \sqrt{(\alpha^{i+1} - \alpha^i)^2 + (\beta^{i+1} - \beta^i)^2}$$