Sample Questions: Maximum Likelihood Part 2

STA312 Spring 2019. Copyright information is at the end of the last page.

1. Let $X_1, \ldots, X_n$ be independent $N(\mu, \sigma^2)$ random variables.

   (a) Derive formulas for the maximum likelihood estimates of μ and σ^2. We will establish that it's a maximum later. Show your work and **circle your final answer**.

$$\ell(\mu,\sigma^2) = \log \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x_i-\mu)^2}$$

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

$$= -\frac{n}{2}\log\sigma^2 - \frac{n}{2}\log(2\pi) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2 \quad (*)$$

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

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

$$= \frac{n}{\sigma^2}(\bar{x}-\mu) \stackrel{set}{=} 0$$

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

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

$$= \frac{-n}{2\sigma^2} + \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^4} \stackrel{set}{=} 0$$

$$\text{Have} \quad \overset{(1)}{\frac{n}{\sigma^2}(\bar{x}-\mu)=0}, \quad \overset{(2)}{\frac{-n}{2\sigma^2}+\frac{\Sigma(x_i-\mu)^2}{2\sigma^4}=0}$$

From (1) $\bar{x}-\mu=0 \Rightarrow \mu=\bar{x}$

From (2) $$\frac{\Sigma(x_i-\bar{x})^2}{\cancel{2}\sigma^4}=\frac{n}{\cancel{2}\sigma^2}$$

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

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

$$\hat{\mu}=\bar{X}, \quad \hat{\sigma}^2=\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{n}$$

(b) Calculate the Hessian of the minus log likelihood function: $\mathbf{H} = \left[\frac{\partial^2(-\ell)}{\partial\theta_i\partial\theta_j}\right]$. Show your work.

Have $\frac{\partial \ell}{\partial \mu} = \frac{n}{\sigma^2}(\bar{x}-\mu) = n(\sigma^2)^{-1}(\bar{x}-\mu)$

And $\frac{\partial \ell}{\partial \sigma^2} = \frac{-n}{2\sigma^2} + \frac{\sum(x_i-\mu)^2}{2\sigma^4}$

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

$$-\frac{\partial^2 \ell}{\partial \mu^2} = \frac{\partial}{\partial \mu}\left(\frac{\mu n}{\sigma^2} - \frac{n\bar{x}}{\sigma^2}\right) = \frac{n}{\sigma^2} - 0$$

$$-\frac{\partial^2 \ell}{\partial {\sigma^2}^2} = \left(+\frac{n}{2}(+1)(\sigma^2)^{-2} + \frac{1}{\cancel{2}}\sum_{i=1}^{n}(x_i-\mu)^2(-\cancel{2})(\sigma^2)^{-3}\right)(-1)$$

$$= \frac{-n}{2\sigma^4} + \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{\sigma^6}$$

$$-\frac{\partial^2 \ell}{\partial \sigma^2 \partial \mu} = (-1)\frac{\partial}{\partial \sigma^2}\left(n(\sigma^2)^{-1}(\bar{x}-\mu)\right)$$

$$= (+1)\, n(\bar{x}-\mu)(-1)(\sigma^2)^{-2}$$

$$= \frac{n(\bar{x}-\mu)}{\sigma^4}$$

$$H = \begin{pmatrix} \dfrac{n}{\sigma^2} & \dfrac{n(\bar{x}-\mu)}{\sigma^4} \\[2ex] \dfrac{n(\bar{x}-\mu)}{\sigma^4} & \dfrac{\sum_{i=1}^{n}(x_i-\mu)^2}{\sigma^6} - \dfrac{n}{2\sigma^4} \end{pmatrix}$$

$$H(\hat{\theta}) = \begin{pmatrix} \dfrac{n}{\hat{\sigma}^2} & 0 \\[2ex] 0 & \dfrac{n\cancel{\hat{\sigma}^2}}{\hat{\sigma}^4\cancel{\hat{\sigma}^2}} - \dfrac{n}{2\hat{\sigma}^4} \end{pmatrix}$$

$$= \begin{pmatrix} \dfrac{n}{\hat{\sigma}^2} & 0 \\[2ex] 0 & \dfrac{n}{2\hat{\sigma}^4} \end{pmatrix}$$

(c) Give $\widehat{\mathbf{V}}_n$, the estimated asymptotic variance-covariance matrix of the MLE. Show some work.

$$\hat{V}_n = H(\hat{\theta})^{-1} = \begin{pmatrix} \frac{\hat{\sigma}^2}{n} & 0 \\ 0 & \frac{2\hat{\sigma}^4}{n} \end{pmatrix}$$

(d) Consider a large-sample Z-test of $H_0 : \mu = \mu_0$. Give an explicit formula for the test statistic. This is something you would be able to compute with a calculator given $\widehat{\mu}$ and $\widehat{\sigma}^2$.

$$Z = \frac{\bar{x} - \mu_0}{\sqrt{\hat{\sigma}^2 / n}} = \frac{\sqrt{n}\,(\bar{x} - \mu)}{\hat{\sigma}}$$

(e) Consider a large-sample Z-test of $H_0 : \sigma^2 = \sigma_0^2$. Give an explicit formula for the test statistic. This is something you would be able to compute with a calculator.

$$Z = \frac{\hat{\sigma}^2 - \hat{\sigma}_0^2}{\sqrt{\frac{2\hat{\sigma}^4}{n}}} = \frac{\sqrt{n}\,(\hat{\sigma}^2 - \sigma_0^2)}{\sqrt{2}\,\hat{\sigma}^2}$$

(f) Consider the large-sample likelihood ratio test of $H_0 : \mu = \mu_0$. Derive an explicit formula for the test statistic G^2. Show your work and *keep simplifying!*.

$$G^2 = -2\log\left(\frac{L(\hat{\theta}_0)}{L(\hat{\theta})}\right) \quad \text{Need } \hat{\theta}_0 = (\hat{\mu}_0, \hat{\sigma}_0^2)$$

$$\hat{\mu}_0 = \mu_0 \quad \text{Find } \hat{\sigma}_0^2$$

Fixed

$$\ell(\mu_0, \sigma^2) = \log\left((\sigma^2)^{-\frac{n}{2}}(2\pi)^{-\frac{n}{2}} \operatorname{Exp}\left\{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu_0)^2(\sigma^2)^{-1}\right\}\right)$$

$$= -\frac{n}{2}\log\sigma^2 - \frac{n}{2}\log 2\pi - \frac{1}{2}\sum_{i=1}^{n}(x_i-\mu_0)^2(\sigma^2)^{-1}$$

$$\ell' = -\frac{n}{2\sigma^2} - 0 - \frac{1}{2}\sum_{i=1}^{n}(x_i-\mu_0)^2(-1)(\sigma^2)^{-2}$$

$$= \frac{-n}{\cancel{2\sigma^2}} + \frac{1}{2}\frac{\sum(x_i-\mu_0)^2}{\sigma^2\sigma^2} \overset{set}{=} 0$$

$$\Rightarrow \frac{n}{\cancel{2\sigma^2}} = \frac{\sum(x_i-\mu_0)^2}{\cancel{2\sigma^2}\sigma^2}$$

$$\Rightarrow \sigma^2 = \frac{\sum(x_i-\mu_0)^2}{n}$$

$$\hat{\sigma}_0^2 = \frac{\sum_{i=1}^{n}(x_i-\mu_0)^2}{n}$$

$$G^2 = -2\log\frac{L(\hat{\theta}_0)}{L(\hat{\theta})}$$

$$= -2\log\frac{(\hat{\sigma}_0^2)^{-\frac{n}{2}}\cancel{(2\pi)^{-\frac{n}{2}}}\operatorname{Exp}\left\{-\frac{1}{2\hat{\sigma}_0^2}\sum_{i=1}^{n}(x_i-\mu_0)^2\right\}}{(\hat{\sigma}^2)^{-\frac{n}{2}}\cancel{(2\pi)^{-\frac{n}{2}}}\operatorname{Exp}\left\{-\frac{1}{2\hat{\sigma}^2}\sum_{i=1}^{n}(x_i-\bar{x})^2\right\}}$$

$$= -2\log\left(\frac{\hat{\sigma}_0^2}{\hat{\sigma}^2}\right)^{-\frac{n}{2}}\cancel{\frac{\operatorname{Exp}\left\{-\frac{1}{2\hat{\sigma}_0^2}n\hat{\sigma}_0^2\right\}}{\operatorname{Exp}\left\{-\frac{1}{2\hat{\sigma}^2}n\hat{\sigma}^2\right\}}}$$

$$= -\cancel{2}\left(-\frac{n}{\cancel{2}}\right)\log\left(\frac{\hat{\sigma}_0^2}{\hat{\sigma}^2}\right)$$

$$= n\left(\log\hat{\sigma}_0^2 - \log\hat{\sigma}^2\right)$$

2. The file `http://www.utstat.toronto.edu/~brunner/data/legal/normal.data.txt` has a random sample from a normal distribution.

   (a) Find the maximum likelihood estimates of $\widehat{\mu}$ and $\widehat{\sigma}^2$ numerically. Compare the answer to your closed-form solution.

   (b) Show that the minus log likelihood is indeed minimized at $(\widehat{\mu}, \widehat{\sigma}^2)$ for this data set.

   (c) Calculate the estimated asymptotic covariance matrix of the MLEs.

   (d) Give a "better" estimated asymptotic covariance matrix based on your closed-form solution.

   (e) Calculate a large-sample 95% confidence interval for σ^2.

   (f) Test $H_0 : \mu = 103$ with a

      i. Z-test.

      ii. Likelihood ratio chi-squared test. Compare the closed-form version.

      iii. Wald chi-squared test.

      Give the test statistic and the p-value for each test.

   (g) The coefficient of variation (used in sample surveys and business statistics) is the standard deviation divided by the mean.

      i. Show that multiplication by a positive constant does not affect the coefficient of variation. This is a paper and pencil calculation.

      ii. Give a numerical point estimate of the coefficient of variation for the normal data of this question. Actually, it's the maximum likelihood estimate, because *the invariance principle of maximum likelihood estimation says that the MLE of a function is that function of the MLE.*

      iii. Using the delta method, give a 95% confidence interval for the coefficient of variation. Start with a paper and pencil calculation of $\dot{g}(\boldsymbol{\theta}) = \left(\frac{\partial g}{\partial \theta_1}, \ldots, \frac{\partial g}{\partial \theta_k}\right)$.

$X \sim ?(\mu, \sigma^2)$, $Y = aX$ $a > 0$ $E(Y) = a\mu$, $Var(Y) = a^2\sigma^2$

$C_x = \frac{\sigma}{\mu}$ & $C_y = \frac{\sqrt{a^2\sigma^2}}{a\mu} = \frac{\cancel{a}\sigma}{\cancel{a}\mu} = C_x$

$$g(\theta) = g(\mu, \sigma^2) = \frac{\sqrt{\sigma^2}}{\mu} = (\sigma^2)^{1/2} \mu^{-1}$$

$$\dot{g}(\theta) = \left(\frac{\partial g}{\partial \theta_1}, \dots, \frac{\partial g}{\partial \theta_k} \right) = \left(\frac{\partial g}{\partial \mu}, \frac{\partial g}{\partial \sigma^2} \right)$$

$$= \left(\sigma (-1) \mu^{-2}, \frac{1}{\mu} \frac{1}{2} (\sigma^2)^{-\frac{1}{2}} \right)$$

$$= \left(\frac{-\sigma}{\mu^2}, \frac{1}{2 \mu \sigma} \right)$$