STAT 231: Statistics

Charles Shen

Spring 2016, University of Waterloo

Formulas and Notes.

Assume all log **are in base** e **unless specified.**
I've tried to use ln for consistency,
but there may be a few inconsistency.

Feel free to email feedback to me at echen902@gmail.com.

Contents

1 Statistical Models and Maximum Likelihood Estimation

Definition. *The **relative likelihood function** is defined as*

$$R(\theta) = \frac{L(\theta)}{L(\hat{\theta})} \quad for\ \theta \in \Omega$$

$Note\ that\ 0 \leq R(\theta) \leq 1\ for\ all\ \theta \in \Omega.$

Definition. *The **log likelihood function** is defined as*

$$l(\theta) = \ln L(\theta) \quad for\ \theta \in \Omega$$

1.1 Likelihood Function for Binomial Distribution

The maximum likelihood estimate of θ is $\bar{\theta} = y/n$.

1.2 Likelihood Function for Poisson Distribution

The value $\theta = \bar{y}$ maximizes $l(\theta)$ and so $\hat{\theta} = \bar{y}$ is the maximum likelihood estimate of θ.

1.3 Likelihood Function for Exponential Distribution

The value $\theta = \bar{y}$ maximizes $l(\theta)$ and so $\hat{\theta} = \bar{y}$ is the maximum likelihood estimate of θ for an Exponential Distribution $\sim Exp(\theta)$.

1.4 Likelihood Function for Gaussian Distribution

The maximum likelihood estimate of θ is $\hat{\theta} = (\hat{\mu}, \hat{\sigma})$, where

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} y_i = \bar{y} \quad \text{and} \quad \hat{\sigma} = \left[\frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2\right]^{1/2}$$

Note that $\hat{\sigma} \neq \sigma$ (sample variance).

1.5 Invariance Property of Maximum Likelihood Estimates

Theorem. *If $\hat{\theta}$ is the maximum likelihood estimate of θ, then $g(\hat{\theta})$ is the maximum likelihood estimate of $g(\theta)$.*

2 Estimation

2.1 Confidence Intervals and Pivotal Quantities

In general, construct a pivot using the estimator, use that to construct coverage interval, estimate it and find the confidence interval.

Theorem. Central Limit Theorem
If n is large, and if $Y_1, \ldots, Y_n$ are drawn from a distribution with mean μ and variance σ^2, then $\bar{Y} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$.

For a Binomial Distribution, the confidence interval is

$$\left[\hat{\pi} \pm z^* \sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}\right]$$

where $\hat{\pi} = \frac{y}{n}$, y is the observed data.
To determine the sample size

$$n \geq \left(\frac{z^*}{MoE}\right)^2 \hat{\pi}(1-\hat{\pi})$$

where MoE is the margin of error.
To be conservative, we usually pick $\hat{\pi} = 0.5$ as it maximizes $\hat{\pi}(1-\hat{\pi})$.

2.2 Chi-Squared Distribution $\sim X_k^2$

The Gamma function is

$$\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y} dy \quad \text{for } \alpha > 0$$

Properties of the Gamma Function:

- $\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$
- $\Gamma(\alpha) = (\alpha-1)!$
- $\Gamma(1/2) = \sqrt{\pi}$

The X_k^2 distribution is a continuous family of distributions on $(0, \infty)$ with probability density function

$$f(x;k) = \frac{1}{2^{k/2}\Gamma(k/2)} x^{(k/2)-1} e^{-x/2} \quad \text{for } x > 0$$

where $k \in \{1, 2, \dots\}$ is a parameter of the distribution.
k is referred to as the "degrees of freedom" (d.f) parameter.

For $X \sim X_k^2$

- $E(X) = k$ and $Var(X) = 2k$
- If $k = 1$, $W = Z^2$ and $Z \sim G(0, 1)$
- If $k = 2$, $W \sim Exp(2)$ $(\theta = 2)$
- If k is large, $W \overset{Appr.}{\sim} N(k, 2k)$
- Let X_{k_1}, X_{k_1} be independent random variables with $X_{k_i} \sim X_{k_i}^2$.
  Then $X_{k_1} + X_{k_2} = X_{k_1+k_2}^2$.

2.3 Student's t Distribution

Student's t distribution has probability density function

$$f(t; k) = c_k \left(1 + \frac{t^2}{k}\right)^{-(k+1)/2} \quad \text{for } t \in \Re \text{ and } k = 1, 2, \dots$$

where the constant c_k is given by

$$c_k = \frac{\Gamma(\frac{k+1}{2})}{\sqrt{k\pi}\Gamma(\frac{k}{2})} \quad k \text{ is the } degrees\ of\ freedom$$

Properties of T:

i) Range of T: $(-\infty, \infty)$

ii) T is symmetric around 0

iii) As $k \uparrow$, $T \to Z$

Theorem. *Suppose $Z \sim G(0, 1)$ and $U \sim X_k^2$ independently. Let*

$$T = \frac{Z}{\sqrt{U/k}}$$

$$\to \frac{\bar{Y} - M}{s/\sqrt{n}} \sim t_{n-1}$$

Then T has ***Student's t distribution with k degrees of freedom***.

2.4 Likelihood-Based Confidence Intervals

Theorem. *A $100p\%$ likelihood interval is an approximate $100q\%$ where $q = 2P(Z \leq \sqrt{-2\ln p}) - 1$ and $Z \sim N(0,1)$.*

Example 2.1. Show that a 1% likelihood interval is an approximate 99.8% confidence interval.
Note that $p = 0.01$

$$\begin{aligned} q &= 2P(Z \leq \sqrt{-2\ln(0.01)}) - 1 \\ &\approx 2P(Z \leq 3.03) - 1 \\ &= 2(0.99878) - 1 \\ &= 0.998 = 99.8\% \end{aligned}$$

Theorem. *If a is a value such that*

$$P = 2P(Z \leq a) - 1 \quad \textit{where } Z \sim N(0,1)$$

then the likelihood interval $\{\theta : R(\theta) \geq e^{-a^2/2}\}$ is an approximate $100p\%$ confidence interval.

2.5 Confidence Intervals for Parameters in the $G(\mu, \sigma)$ Model

If $Y_1, \ldots, Y_n$ are independent $N(\mu, \sigma^2)$, then

$$\frac{\bar{Y} - \mu}{s/\sqrt{n}} \sim t_{n-1} \tag{1}$$

$$\frac{(n-1)S^2}{\sigma^2} \sim X^2_{n-1} \tag{2}$$

General Rule:
The Confidence Interval for μ if σ is unknown is

$$\left[\ \bar{y} \pm t^* \frac{s}{\sqrt{n}}\ \right]$$

When σ is unknown, we replace σ by its estimate s, and we use t-pivot.
Confidence interval when σ is known is

$$\left[\ \bar{y} \pm z^* \frac{\sigma}{\sqrt{n}}\ \right]$$

When σ is known, we use z-pivot.

If n is really large, then the t^* value converges to the corresponding z^* value (by Central Limit Theorem).

Confidence Intervals for σ^2 and σ

Theorem. *Suppose $Y_1, Y_2, \ldots, Y_n$ is a random sample from the $G(\mu, \sigma)$ distribution with sample variance S^2. Then the random variable*

$$\frac{(n-1)S^2}{\sigma^2} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(Y_i - \bar{Y})^2$$

has a Chi-squared distribution with $n-1$ degrees of freedom.

Using the theorem, we can construct a $100p\%$ confidence interval for the parameter σ^2 or σ.
Recall this is the same as the equation (2) in this sub-section.
We can find constants a and b such that

$$P(a \leq U \leq b) = p$$

where $U \sim X^2_{n-1}$.
So a $100p\%$ confidence interval for σ^2 is

$$\left[\frac{(n-1)s^2}{b}, \frac{(n-1)s^2}{a}\right]$$

and a $100p\%$ confidence interval for σ is

$$\left[\sqrt{\frac{(n-1)s^2}{b}}, \sqrt{\frac{(n-1)s^2}{a}}\right]$$

Unlike confidence interval for μ, the confidence interval for σ^2 is *not symmetric* about s^2. the estimator of σ^2. The X^2_{n-1} distribution is not a symmetric distribution.

Prediction Interval for a Future Observation
Suppose that $Y \sim G(\mu, \sigma)$, then

$$Y - \widetilde{\mu} = Y - \bar{Y} \sim N\left(0, \sigma^2\left(1 + \frac{1}{n}\right)\right)$$

Also

$$\frac{Y - \bar{Y}}{S\sqrt{1 + \frac{1}{n}}} \sim t_{n-1}$$

is a pivotal quantity which can be used to obtain an interval of values for Y. Let a be a value such that $P(-a \leq T \leq a) = p$ or $P(T \leq a) = (1+p)/2$ which is obtained from tables. Thus

$$\left[\ \bar{y} \pm as\sqrt{1 + \frac{1}{n}}\ \right]$$