Lecture Summary

Recap

Steps to MLE estimation:

- Find the likelihood function.
- Find the log likelihood function.
- Take the derivative to find the global optimum $\hat{\theta}$
- Use the second derivative to check that $\hat{\theta}$ is a maximizer.

Examples

- Let $X_1, \dots, X_n$ be i.i.d. Uniform $[0, \theta]$, where $\theta > 0$. Find $\hat{\theta}$
- Let $X_1, \dots, X_n$ be i.i.d. Uniform $[\theta, \theta + 1]$. Find $\hat{\theta}$

Examples

- Let $X_1, \ldots, X_n$ be i.i.d. Uniform $[0, \theta]$, where $\theta > 0$. Find $\hat{\theta}$
- Let $X_1, \ldots, X_n$ be i.i.d. Uniform $[\theta, \theta + 1]$. Find $\hat{\theta}$
- Does not always exist.
- Is not always unique.

Example

Sampling from a Gamma Distribution.

Suppose that $X_1, \ldots, X_n$ form a random sample from the gamma distribution for which the p.d.f. is as follows:

$$f(x \mid \alpha) = \frac{1}{\Gamma(\alpha)} x^{\alpha-1} e^{-x} \text{ for } x > 0.$$

Find the MLE for α, $(\alpha > 0)$.

Numerical computation

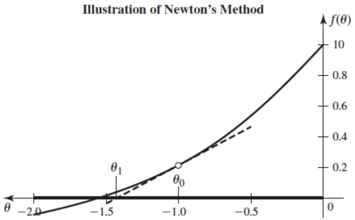

Figure: Enter Caption

Newton's Method.

Let $f(\theta)$ be a real-valued function of a real variable, and suppose that we wish to solve the equation $f(\theta) = 0$. Let θ_0 be an initial guess at the solution. Newton's method replaces the initial guess with the updated guess

$$\theta_1 = \theta_0 - \frac{f(\theta_0)}{}$$

Method of Moments

- $X_1, \ldots, X_n \sim$ Distribution(k-dimensional parameter θ) with at least k finite moments.
- Reminder: j-th moment of X_i: $\mu_j(\theta) = E\left(X_i^j \mid \theta\right)$
- Define the sample j-th moment: $m_j = \frac{1}{n}\sum_{i=1}^{n} X_i^j$ for $j = 1, \ldots, k$.
- Set up the k equations $m_j = \mu_j(\theta)$ and then solve for θ.

Reminder

Definition (Gamma distribution)

A RV X has the Gamma distribution with parameters $\alpha, \beta > 0$ if

$$f(x|\alpha,\beta) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} & x > 0 \\ 0 & otherwise \end{cases}$$

- Suitable for RVs in $(0, \infty)$
- Parameter space: $\alpha, \beta > 0$.
- $E(X) = \frac{\alpha}{\beta}, Var(X) = \frac{\alpha}{\beta^2}$.
- MGF: $\left(1 - \frac{t}{\beta}\right)^{-\alpha}$ for $t < \beta$

Summary - Estimators

Bayesian Inference

- Mean of the posterior distribution is a Bayes estimator for squared error loss.
- Median of the posterior is a Bayes estimator for absolute loss.
- Mode of the posterior is the Maximum A Posteriori (MAP) estimator (also a popular choice).

MLE

- MLE estimator maximizes the likelihood function.

MOM

- Method of moments estimator matches moments to sample moments.

Invariance Property of MLE

If $\hat{\theta}$ is the maximum likelihood estimator of θ and if g is a one-to-one function, then $g(\hat{\theta})$ is the maximum likelihood estimator of $g(\theta)$.

Example

Assume $X_1, \ldots, X_n \sim Expo(\theta)$. θ is interpreted as the failure rate of electronic components.

- Find the MLE estimate for the failure rate $\hat{\theta}_{MLE}$.
- Find the MLE estimate for the average lifetime $\psi = 1/\theta$.

Properties of Estimators

An estimator is a function of the random sample:

$$\hat{\theta}(X_1, \ldots, X_n) = f(X_1, \ldots, X_n)$$

Consistent Estimators

A sequence of estimators that converges in probability to the unknown value of the parameter being estimated, as $n \to \infty$, is called a consistent sequence of estimators.

Asymptotically Normal Estimators

A sequence of estimators that converges in distribution to a normal distribution is called an asymptotically normal sequence of estimators.

MLE estimators are usually consistent and asymptotically normal

Unbiased Estimators

Unbiased Estimator/Bias.

An estimator $\delta(\boldsymbol{X})$ is an unbiased estimator of a function $g(\theta)$ of the parameter θ if $E_\theta[\delta(\boldsymbol{X})] = g(\theta)$ for every possible value of θ. An estimator that is not unbiased is called a biased estimator. The difference between the expectation of an estimator and $g(\theta)$ is called the bias of the estimator. That is, the bias of δ as an estimator of $g(\theta)$ is $E_\theta[\delta(\boldsymbol{X})] - g(\theta)$, and δ is unbiased if and only if the bias is 0 for all θ.