7.2 Maximum Likelihood Estimators

- Let X_i's be independent exponential random variables each having the same unknown mean θ.

  Then the joint probability distribution function of X_i's is,

  $$f(x_1, x_2, \ldots, x_n) = f_{X_1}(x_1)\, f_{X_2}(x_2) \cdots f_{X_n}(x_n) = \frac{1}{\theta}e^{-\frac{x_1}{\theta}} \cdot \frac{1}{\theta}e^{-\frac{x_2}{\theta}} \cdots \frac{1}{\theta}e^{-\frac{x_n}{\theta}}$$

  $$= \frac{1}{\theta^n} e^{-\sum_{i}^{n} \frac{x_i}{\theta}}$$, is the function of θ. $\underbrace{f(x_1, \ldots x_n \mid \theta)}_{\text{likelihood function of } \theta}$

  $\Rightarrow$ the maximum likelihood estimate $\hat{\theta}$ is defined to be that value of θ maximizing $f(x_1, \ldots, x_n \mid \theta)$.

  $\Rightarrow$ In determining the maximizing value of θ, it is often useful to use the fact that $f(x_1, \ldots, x_n \mid \theta)$ and $\log[f(x_1, \ldots, x_n \mid \theta)]$ have their maximum at the same value of θ. Hence, we may also obtain $\hat{\theta}$ by maximizing $\log[f(x_1, \ldots, x_n \mid \theta)]$

7.3 Interval Estimates

- Although $\bar{X}$ is the maximum likelihood estimator for μ, we can only expect $\bar{X}$ to be close to μ. So, rather than a point estimate, it is sometimes more valuable to be able to specify an interval.
- Since the point estimator $\bar{X}$ is normal with mean μ and variance $\frac{\sigma^2}{n}$,

  $$P\left\{\bar{X} - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right\} = 1 - \alpha$$

7.3.1 Confidence Interval For A Normal Mean When The Variance Is Unknown

- it still follows that $\sqrt{n}\frac{(\bar{X} - \mu)}{S}$ is a t-random variable with $n-1$ degrees of freedom

  $$P\left\{\bar{X} - t_{\alpha/2, n-1}\frac{S}{\sqrt{n}} < \mu < \bar{X} + t_{\alpha/2, n-1}\frac{S}{\sqrt{n}}\right\} = 1 - \alpha$$

7.3.2 Prediction Intervals

- Suppose $\bar{X}$ is normal with mean μ and variance $\frac{\sigma^2}{n}$, and X_{n+1} is normal with mean μ and variance σ^2, then it follows that $X_{n+1} - \bar{X}_n$ is normal with mean 0 and variance $\frac{\sigma^2}{n} + \sigma^2$.

  $\Rightarrow$ $\frac{X_{n+1} - \bar{X}_n}{\sigma\sqrt{1 + 1/n}}$ is a standard normal random variable.
- Replacing σ by its estimator S_n will yield a t-random variable with $n-1$ degrees of freedom.

  $\Rightarrow$ $\frac{X_{n+1} - \bar{X}_n}{S_n\sqrt{1 + 1/n}}$ is a t-random variable with $n-1$ degrees of freedom

7.3.3 Confidence Intervals For The Variance of a Normal Distribution

- If the samples are from a normal distribution, the confidence interval for σ^2 is found by using

  $$(n-1)\frac{S^2}{\sigma^2} \sim \chi^2_{n-1} \Rightarrow P\left\{\chi^2_{1-\alpha/2, n-1} \le (n-1)\frac{S^2}{\sigma^2} \le \chi^2_{\alpha/2, n-1}\right\} = 1 - \alpha$$

  $$\Rightarrow P\left\{\frac{(n-1)S^2}{\chi^2_{\alpha/2, n-1}} \le \sigma^2 \le \frac{(n-1)S^2}{\chi^2_{1-\alpha/2, n-1}}\right\} = 1 - \alpha$$

7.4 ## Estimating the Difference in Means of Two Normal Populations

- Let $X_1, X_2, \ldots, X_n$ be a sample of size n from a normal population having mean μ_1 and variance σ_1^2, and let $Y_1, Y_2, \ldots, Y_m$ be a sample of size m from a different normal population having mean μ_2 and variance σ_2^2, and suppose they are independent, and $\mu_1 - \mu_2$ is of interest.

$$\Rightarrow \bar{X} - \bar{Y} \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}\right)$$

$$\Rightarrow \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}} \sim N(0,1)$$

$$\Rightarrow P\left\{\bar{X} - \bar{Y} - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}} < \mu_1 - \mu_2 < \bar{X} - \bar{Y} + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}\right\} = 1 - \alpha$$

- However, to utilize the foregoing to obtain a confidence interval, we need its distribution and it must not depend on any of the unknown parameters σ_1^2 and σ_2^2. So, suppose that there is a common variance σ^2.

$$\Rightarrow (n-1)\frac{S_1^2}{\sigma^2} \sim \chi^2_{n-1} \quad , \quad (m-1)\frac{S_2^2}{\sigma^2} \sim \chi^2_{m-1}$$

$$\Rightarrow (n-1)\frac{S_1^2}{\sigma^2} + (m-1)\frac{S_2^2}{\sigma^2} \sim \chi^2_{n+m-2}$$

$$\Rightarrow S^2_{pooled} = \frac{(n-1)S_1^2 + (m-1)S_2^2}{n+m-2}$$

$$\Rightarrow P\left\{\bar{X} - \bar{Y} - t_{\alpha/2, n+m-2} \cdot S_p\sqrt{\frac{1}{n} + \frac{1}{m}} \le \mu_1 - \mu_2 \le \bar{X} - \bar{Y} + t_{\alpha/2, n+m-2} \cdot S_p\sqrt{\frac{1}{n} + \frac{1}{m}}\right\} = 1 - \alpha$$

7.5 ## Approximate Confidence Interval for the Mean of a Bernoulli Random Variable

$$\frac{X - np}{\sqrt{n\hat{p}(1-\hat{p})}} \sim N(0,1) \Rightarrow P\left\{\hat{p} - z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n} < p < \hat{p} + z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n}\right\} \approx 1 - \alpha$$

- If we were to specify an approximate $100(1-\alpha)\%$ CI for p no greater than some given length b. Note that the length of the approximate $100(1-\alpha)\%$ CI for p is $2 \cdot z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n} \approx b$.

$$\Rightarrow (2z_{\alpha/2})^2 p(1-p)/n = b^2$$

$$\Rightarrow n = \frac{(2z_{\alpha/2})^2 p^*(1-p^*)}{b^2}$$

, That is, if k items were initially sampled to obtain the preliminary estimate of p, then an additional $n-k$ items should be sampled.

Type of Interval	Confidence Interval
Two-sided	$\hat{p} \pm z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n}$
One-sided lower	$\left(-\infty, \hat{p} + z_{\alpha}\sqrt{\hat{p}(1-\hat{p})/n}\right)$
One-sided upper	$\left(\hat{p} - z_{\alpha}\sqrt{\hat{p}(1-\hat{p})/n}, \infty\right)$

7.6 Confidence Interval of the Mean of the Exponential Distribution

$$\frac{2}{\theta}\sum_{i=1}^{n} X_i \sim \chi^2_{2n} \Rightarrow P\left\{\chi^2_{1-\alpha/2,2n} < \frac{2}{\theta}\sum_{i=1}^{n} X_i < \chi^2_{\alpha/2,2n}\right\} = 1-\alpha$$

$$\Rightarrow P\left\{\frac{2\sum_{i=1}^{n} X_i}{\chi^2_{\alpha/2,2n}} < \theta < \frac{2\sum_{i=1}^{n} X_i}{\chi^2_{1-\alpha/2,2n}}\right\} = 1-\alpha$$

7.7 Evaluating a Point Estimator

Definition of Bias