Statistics (Math 324)

Lecture 5: Thursday Jan. 25th, 2018

Confidence Intervals

- Random interval:

  An interval whose end point(s) are r.v.s is called a random interval.

- Confidence interval:

  A 100(1-α)% confidence interval for a parameter θ is a random interval $(\hat{\theta}_L(X), \hat{\theta}_U(X))$ such that

$$P(\hat{\theta}_L(X) < \theta < \hat{\theta}_U(X)) = 1-\alpha$$

- Pivotal Quantity:

  A function of the observations $X_1, \ldots, X_n$ and some unknown parameters, ideally just the parameter(s) of interest, whose distribution DOES NOT depend on any unknown parameter is called a pivotal quantity.

Pivotal Quantities play a central role in theory confidence Intervals.

Example: Let $X_i \overset{iid}{\sim} N(\mu, \sigma^2)$, $i=1,2,\ldots,n$ where σ^2 is known, but μ is unknown. We show that $\bar{X}_n \sim N(\mu, \frac{\sigma^2}{n})$ where $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$

Recall that there are three methods for finding the distribution of a function of random variables

- Method of Transformation:

This is essentially Theorem of Change of variables in Calculus.

- Method of distribution:

On this method we connect the cdf of the new variable to the cdf of the original variables.

Example: suppose $X_i \overset{iid}{\sim} f$, $i=1,\dots,n$ are continuous r.v.s with pdf f and cdf F. Define $X_{(n)} = \max\limits_{1\le i\le n} X_i$. Then

$$F_{X_{(n)}}(t) = P(X_{(n)} \le t) = P(X_1 \le t, \dots, X_n \le t)$$

$$\underset{i=1}{\overset{n}{\perp\!\!\!\perp}} X_i \Rightarrow \quad = \prod_{i=1}^n P(X_i \le t) = \prod_{i=1}^n F_{X_i}(t)$$

$$\text{Identically distributed} \quad = \prod_{i=1}^n F(t) = F^n(t)$$

Thus

$$f_{X_{(n)}}(t) = \frac{d}{dt} F_{X_{(n)}}(t) = \frac{d}{dt} F^n(t) = n f(t) F^{n-1}(t)$$

- Method of moment generating function (mgf)

This method is essentially based on the mgf. In this method we try to connect the mgf of the new variable to the mgf of the original variables.

Ex. suppose $X_i \sim N(\mu_i, \sigma_i^2)$, $i=1,2,\dots,n$. suppose that X_is are independent. Define $S = \sum_{i=1}^n X_i$.

Then

$$m_S(t) = E[e^{tS}] = E[e^{t\sum_{i=1}^n X_i}] = E[\prod_{i=1}^n e^{tX_i}]$$

$$\text{Using independence } (\perp\!\!\!\perp) = \prod_{i=1}^n E[e^{tX_i}] = \prod_{i=1}^n m_{X_i}(t) = \prod_{i=1}^n e^{\mu_i t + \frac{\sigma_i^2 t^2}{2}}$$

$$= \exp\{t\sum_{i=1}^n \mu_i + \frac{t^2}{2}\sum_{i=1}^n \sigma_i^2\}.$$

Thus $\boxed{S \sim N(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2)}$

If we further assume that X_i's are identically distributed then $\mu_i = \mu$, $\sigma_i^2 = \sigma$, $i=1,2,\dots,n$ and therefore we have

$m_S(t) = \exp\{n\mu t + \frac{n\sigma^2 t^2}{2}\}$ and hence

$$\boxed{S \sim N(n\mu, n\sigma^2)}$$

Then $m_{\bar{X}_n}(t) = E[e^{t\bar{X}_n}] = E[e^{t \cdot \frac{1}{n}\sum_{i=1}^n X_i}]$

$t^* = \frac{t}{n} \Rightarrow \quad = E[e^{t^* S}] = m_S(t^*) = e^{n\mu t^* + \frac{n\sigma^2 t^{*2}}{2}}$

and therefore $= \exp\{\mu t + \frac{(\frac{\sigma^2}{n})t^2}{2}\}$

$$\bar{X}_n \sim N(\mu, \frac{\sigma^2}{n}). \quad (1)$$

Note further that if $X \sim N(\mu, \sigma^2)$, then

$Z = \frac{X-\mu}{\sigma} \sim N(0,1)$. We prove a general form of this. Let $X \sim N(\mu, \sigma^2)$. Then

$$aX + b \sim N(a\mu + b, a^2\sigma^2)$$

for any a, b constant. Let $V = aX + b$. Then

$m_V(t) = E[e^{tV}] = E[e^{t(aX+b)}] = E[e^{taX + tb}]$

$= E[\underbrace{e^{tb}}_{\text{constant}} \cdot e^{\overbrace{ta}^{t^*}X}] = e^{tb} \cdot E[e^{t^*X}]$

$= e^{tb} \cdot m_X(t^*) = e^{tb} \cdot e^{\mu t^* + \frac{\sigma^2 t^{*2}}{2}}$

$= e^{tb} \cdot e^{\mu t a + \frac{\sigma^2 t^2 a^2}{2}}$

$= \exp\{t(a\mu + b) + \frac{(a^2\sigma^2)t^2}{2}\}$

Thus $aX + b \sim N(a\mu + b, a^2\sigma^2)$. Now

$Z = \frac{X-\mu}{\sigma} = \frac{X}{\sigma} - \frac{\mu}{\sigma} = \frac{1}{\sigma}X - \frac{\mu}{\sigma}$

$= aX + b$ where $a = \frac{1}{\sigma}$ and $b = -\frac{\mu}{\sigma}$

Hence $Z \sim N\left(\underbrace{\frac{1}{\sigma}\mu + \left(-\frac{\mu}{\sigma}\right)}_{0}, \underbrace{\left(\frac{1}{\sigma}\right)^2\sigma^2}_{1}\right)$ and therefore

$$Z \sim N(0,1). \quad (2)$$

Using (1) and (2)

$$\frac{\bar{X}_n - \mu}{\sqrt{\frac{\sigma^2}{n}}} = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \sim N(0,1).$$ This means that

$$\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}$$ is a pivotal quantity.

To summarize:

$$X_i \overset{iid}{\sim} N(\mu, \sigma^2) \Rightarrow \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}$$ is pivotal Quantity

Notice that using the table for Normal dist.

$$P\left(\left|\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}\right| \leq 1.96\right) = 0.95$$ or

equivalently

$$P\left(\bar{X}_n - 1.96\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X}_n + 1.96\frac{\sigma}{\sqrt{n}}\right) = 0.95$$

This then means that

$$\left(\bar{X}_n - 1.96\frac{\sigma}{\sqrt{n}}, \bar{X}_n + 1.96\frac{\sigma}{\sqrt{n}}\right)$$

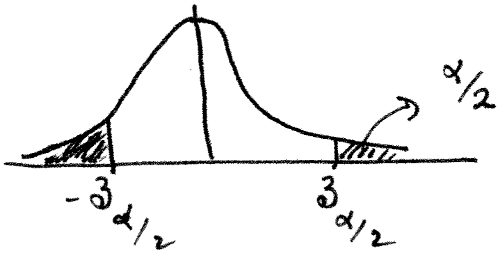
covers the true μ with 95% probability.

Thus a $100(1-\alpha)\%$ confidence interval for μ when $X_i \overset{iid}{\sim} N(\mu, \sigma^2)$ and σ^2 is known is

$$\bar{X}_n \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

where

$P(Z > z_{\alpha/2}) = \alpha/2$, $Z \sim N(0,1)$



<u>Remark</u>: In real applications we compute $\bar{X}_n$ and obtain an interval, say $(125, 135)$. Now either this interval covers the true μ or it does not. Then the question is what do we mean by a 95% C.I.?

Note that the $100(1-\alpha)\%$ confidence is the property of the procedure. It means that out of the all possible intervals of the form $(\bar{X}_n \pm 1.96 \frac{\sigma}{\sqrt{n}})$ that we can make by taking samples of size n from $N(\mu, \sigma^2)$, 95% of them cover the true μ. Now in a real application when we make one of such intervals by taking a random sample of size n from $N(\mu, \sigma^2)$, it is like taking one of those intervals randomly. Since that 95% of them cover μ, my chance of selecting an interval that covers μ is 95%. Thus I can take a bet 19 to 1 that the interval I select covers μ.

- Large Sample Confidence Interval:

The derivation of the pivotal quantity in the above example totally hinges over the normality assumption, i.e. $X_i \sim N(\mu, \sigma^2)$. What happens if we do not know the parametric form of the population distribution?

— Central Limit Theorem (CLT)
(The baby form)

Suppose $X_1, \ldots, X_n$ are independent random variables with common mean μ and variance σ^2. Then

$$\frac{\bar{X}_n - \mu}{\left(\frac{\sigma}{\sqrt{n}}\right)} \overset{app}{\sim} N(0,1)$$

when n is large enough.

This powerful theorem then implies that $\frac{\bar{X}_n - \mu}{\left(\frac{\sigma}{\sqrt{n}}\right)}$ is approximately a pivotal quantity distributed according to $N(0,1)$ for large enough n regardless of population distribution provided that the conditions of the CLT are met.

σ^2 known $\rightarrow$ $\left(\bar{X}_n \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right)$ is a $100(1-\alpha)\%$ C.I for μ

σ^2 unknown $\rightarrow$ $\left(\bar{X}_n \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right)$ is still a $100(1-\alpha)$ C.I for μ, but not useful.

We need to somehow get rid of the nuisance parameter σ. We can replace σ by S_n where

$$S_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2$$

justification: intuitive; formal

Intuitive : S_n^2 is the sample counterpart, almost, of σ^2. Thus as n increases greater portion of the population and hence our sample gets closer to the population.

Formal! The formal proof comprises three steps:

1- CLT of $\dfrac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}$

2. Consistency of S_n^2 for σ^2, i.e $S_n^2 \xrightarrow{P} \sigma^2$ which we learn in Ch 9. We then use a Thm called continuous mapping theorem which says that if $S_n^2 \xrightarrow{P} \sigma^2$, then $g(S_n^2) \xrightarrow{P} g(\sigma^2)$ for any continuous function. Considering $g(x) = \sqrt{x}$ we obtain $S_n \xrightarrow{P} \sigma$ and hence $\sigma/S \xrightarrow{P} 1$.

3. Cramer's Theorem: This result says that if $V_n \xrightarrow{D} X$ and $Y_n \xrightarrow{P} 1$, then $Y_n \cdot V_n \xrightarrow{D} X$.

$$\frac{\bar{X}_n - \mu}{\frac{S}{\sqrt{n}}} = \underbrace{\frac{\sigma}{S_n}}_{Y_n} \cdot \underbrace{\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}}_{V_n}$$

Note that CLT implies that $V_n \xrightarrow{D} Z$, i.e.

$$\underbrace{F_{V_n}(t)}_{\text{cdf of } V_n} \to F_Z(t) = \underbrace{\int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx}_{\Phi(t),\ \text{cdf of } N(0,1)}$$

Using step 2, $Y_n = \frac{\sigma}{S_n} \xrightarrow{P} 1$ and an application of Cramer's Thm completes the proof.

To summarize

σ^2 known $\to$ $\left(\bar{X}_n \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right)$ is a $100(1-\alpha)\%$ C.I. for μ

σ^2 unknown $\to$ $\left(\bar{X}_n \pm z_{\alpha/2} \frac{S}{\sqrt{n}}\right)$ is a $100(1-\alpha)$ C.I. for μ

To be more precise, these confidence intervals are <u>approximate</u> $100(1-\alpha)\%$ confidence intervals for μ when n is large enough.

So far we focused on C.I. for the population mean. How can we make C.I. for other estimands?

A common, perhaps the most common, method of estimation that we will learn about in Ch 9. is <u>the method of maximum likelihood</u>. Suppose θ is a parameter of interest. Suppose $\hat{\theta}_n = \hat{\theta}(X_1, \ldots, X_n)$ is the maximum likelihood estimate (MLE) of θ based on $X_1, \ldots, X_n$. Then relatively general conditio[ns] we have

$$\frac{\hat{\theta}_n - \theta}{\sqrt{Var(\hat{\theta}_n)}} \overset{app}{\sim} N(0,1) \quad (\star)$$

when n is large enough. We therefore have a general receipe for confidence interval when the sa[mple] size n is large enough, namely

$$\hat{\theta}_n \pm z_{\alpha/2} \sqrt{Var(\hat{\theta}_n)} \quad (\dagger)$$

that is a $100(1-\alpha)\%$ C.I. for θ.

Examples:

1 - $X_i \overset{iid}{\sim} N(\mu, \sigma^2)$, $i = 1, 2, \ldots, n$ ($\sigma^2 \rightarrow$ Known)

We show in Ch 9 that $\bar{X}_n$ is the MLE of μ. Note that $Var(\bar{X}_n) = \frac{\sigma^2}{n}$. Then using $(\dagger)$

$$\bar{X}_n \pm z_{\alpha/2} \sqrt{\frac{\sigma^2}{n}}$$ is a $100(1-\alpha)\%$ C.I. for μ

2 - $X_i \overset{iid}{\underset{i=1,\dots,n}{\sim}} \text{Bernoulli}(p)$, i.e.

$$X_i = \begin{cases} 1 & p \\ 0 & 1-p \end{cases}$$

Then $\hat{p}_n = \frac{1}{n}\sum_{i=1}^{n} x_i$ is the MLE of p. Thus using (†)

$$\hat{p}_n \pm z_{\alpha/2}\sqrt{Var(\hat{p}_n)}$$ is a $100(1-\alpha)\%$ C.I. for p.

Note that $Var(\hat{p}_n) = \frac{p(1-p)}{n}$. We have two choices

- replace p by $\hat{p}_n$ in $Var(\hat{p}_n)$
- replace $p(1-p)$ in $Var(\hat{p}_n)$ by $\frac{1}{4}$ to find a conservatively large C.I. for p

i.e.

- $$\hat{p}_n \pm z_{\alpha/2}\frac{\sqrt{\hat{p}_n(1-\hat{p}_n)}}{\sqrt{n}}$$
- $$\hat{p}_n \pm z_{\alpha/2}\cdot\frac{1}{2\sqrt{n}}$$

3 - Now suppose $X_i \overset{iid}{\sim} Ber(p)$, $i=1,\dots,n$ and we are interested in $\theta = p(1-p)$, the variance. An interesting property of MLE is the invariance, i.e. if $\hat{\theta}_n$ is the MLE of θ, then $h(\hat{\theta}_n)$ is the MLE of $h(\theta)$. The invariance property then implies that $\hat{\theta}_n = \hat{p}_n(1-\hat{p}_n)$ is the MLE of $p(1-p)=\theta$

The $100(1-\alpha)\%$ C.I for $\theta = p(1-p)$ is

$$\hat{\theta}_n \pm z_{\alpha/2} \sqrt{Var(\hat{\theta}_n)} \; .$$

- Small Sample Confidence Intervals

Unlike the large sample case, there is no general receipe like (*) using which we can find an approximate pivotal quantity. In fact, there is on the paper, but only gives fruits in special cases. To summarize, small sample problems are solved mostly case by case. A case of particular importance is the Normal case. We will learn about the importance of this case when we discuss regression and ANOVA (Analysis of Variance).

- Normal case.

suppose $X_i \overset{iid}{\sim} N(\mu, \sigma^2)$, $i = 1, 2, -, n$

nuisance

of interest

where n, the sample size, is NOT large.

We learnt that when $X_i \overset{iid}{\sim} N(\mu, \sigma^2)$, $i = 1, -, n$

$$\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \overset{Exact}{\sim} N(0,1) \qquad (\ddagger)$$

This by itself is not, of course, useful since σ is NOT known. We discussed in previous section at length why we can replace σ by S when n is large enough. The formal justification is NOT applicable now since n is small, the intuitive justification still stands though.

Replacing σ by S in (‡) changes the picture a li
Given that S has the same spirit as σ, though in a smaller scale the dist. of $T = \frac{\bar{X}_n - \mu}{\frac{S}{\sqrt{n}}}$ still has a bell curve shape. The tails of the dist., however, die out much more slowly than those of normal dist. Heavier tails mean much more variability and this should perhaps be expected since by replacing σ by S which can be crude estimate when n is small, can add quite a bit to the variability. This is, of course, an intuitive argument. Following we present the sketch of a formal argument.

- step 1.

$$X_i \overset{iid}{\sim} N(\mu, \sigma^2) \Rightarrow \bar{X}_n \sim N(\mu, \sigma^2/n)$$

$$\Rightarrow \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

- step 2.

$$X_i \overset{iid}{\sim} N(\mu, \sigma^2) \Rightarrow \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$$

proof:

$$\sum_{i=1}^{n} (X_i - \mu)^2 = \sum_{i=1}^{n} [(X_i - \bar{X}_n) + (\bar{X}_n - \mu)]^2$$

$$= \sum_{i=1}^{n} (X_i - \bar{X}_n)^2 + n(\bar{X}_n - \mu)^2 + 2(\bar{X}_n - \mu)\underbrace{\sum_{i=1}^{n}(X_i - \bar{X}_n)}_{0}$$

$$= (n-1)S^2 + n(\bar{X}_n - \mu)^2$$

Dividing both sides by σ^2 we obtain

$$\underbrace{\sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma}\right)^2}_{W} = \underbrace{\frac{(n-1)S^2}{\sigma^2}}_{U} + \underbrace{\left(\frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}}\right)^2}_{V} \quad (28)$$

Now note that

$$X_i \overset{iid}{\sim} N(\mu, \sigma^2) \Rightarrow \frac{X_i - \mu}{\sigma} \sim N(0,1) \Rightarrow$$

$$\left(\frac{X_i - \mu}{\sigma}\right)^2 \sim \chi^2_1 \Rightarrow \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma}\right)^2 \sim \chi^2_n$$

(Exercise). This Thm 7.2, p 356.

Thus $W \sim \chi^2_n$. On the other hand, using step:

$$\left(\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}\right)^2 \sim \chi^2_1$$

- step 3:

If $X_i \overset{iid}{\sim} N(\mu, \sigma^2)$, then $\bar{X}_n \perp\!\!\!\perp S^2$

- step 4:

$$m_W(t) = E[e^{tW}] = E[e^{t(U+V)}]$$

$$= E[e^{tU} \cdot e^{tV}]$$

$U \perp\!\!\!\perp V$ using step 3 $\quad = E[e^{tU}] \cdot E[e^{tV}] = m_U(t) \cdot m_V(t)$

Thus $m_U(t) = \dfrac{m_W(t)}{m_V(t)} = \dfrac{(1-2t)^{-n/2}}{(1-2t)^{-1/2}} = (1-2t)^{-\frac{n-1}{2}}$

which implies that $U \sim \chi^2_{(n-1)}$

- step 5:

If $Z \sim N(0,1)$, $U \sim \chi^2_{\nu}$ and $Z \perp\!\!\!\perp U$, then

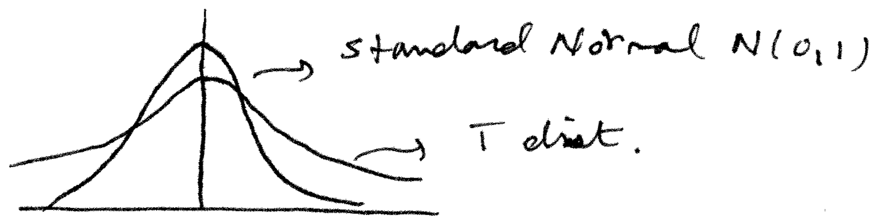
$$\frac{Z}{\sqrt{\frac{U}{\nu}}} \sim T_{n-1} \quad (\text{Ex. } 7.30, \ p\, 367)$$

- Step 6:

$$T_{n-1} = \frac{\bar{X}_n - \mu}{\frac{s}{\sqrt{n}}} = \frac{\left(\frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}}\right)}{\sqrt{\frac{\left[\frac{(n-1)s^2}{\sigma^2}\right]}{(n-1)}}} = \frac{Z}{\sqrt{\frac{U}{v}}}$$

The pdf of T_v is

$$f_{T_v}(t) = \frac{\Gamma\left[\frac{v+1}{2}\right]}{\Gamma\left(\frac{v}{2}\right)\sqrt{v\pi}} \left(1 + \frac{t^2}{v}\right)^{-\frac{v+1}{2}} \qquad -\infty < t < +\infty$$



$$E[T_v^r] = \begin{cases} 0 & \text{if } r < v \text{ and } r \text{ is odd} \\ v^{\frac{r}{2}} \cdot \dfrac{\Gamma\left(\frac{r+1}{2}\right)\Gamma\left(\frac{v-r}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{v}{2}\right)} & \text{if } r < v \text{ and } r \text{ is even} \end{cases}$$

Thus, if $X_i \overset{iid}{\sim} N(\mu, \sigma^2)$, $i = 1, \ldots, n$, μ & σ^2 both <u>unknown</u>

$$\bar{X}_n \pm t_{(n-1), \alpha/2} \frac{s}{\sqrt{n}}$$

provides a $100(1-\alpha)\%$ C.I. for μ, where

$$P\left(T_{(n-1)} > t_{(n-1), \alpha/2}\right) = \alpha/2$$

- Pivotal quantity and probability Integral Transform

Suppose X is a continuous r.v. with pdf f and cdf F. Then $F(X) \sim Unif(0,1)$ (Exercise).

This result is referred to as the Probability Integral transform. Now suppose $X_i \overset{iid}{\sim} F$. Then

$$F(X_i) \sim \text{Unif}(0,1) \Rightarrow -2\ln F(X_i) \sim \chi^2_2$$

$$\Rightarrow -2\sum_{i=1}^{n} \ln F(X_i) \sim \chi^2_{2n} \quad (\text{Exercise})$$

$$\Rightarrow -2\sum_{i=1}^{n} \ln[1-F(X_i)] \sim \chi^2_{2n}$$

There is hence a general receipe for finding a pivot quantity when we have samples from continuous r.v.s. The usefulness of this pivotal quantity depends on the form of F, the cdf of X.

Suppose $X_i \overset{iid}{\sim} Exp(\lambda)$, $i=1,2,\dots,n$, i.e.

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x>0 \\ 0, & \text{o.w.} \end{cases}$$

Then

$$F(x) = \int_0^x f(t)\,dt = 1-e^{-\lambda x}, \quad x>0$$

$$F(x) = \begin{cases} 0 & \text{if } x \le 0 \\ 1-e^{-\lambda x} & \text{if } x>0 \end{cases}$$

Using the above discussion

$$-2\sum_{i=1}^{n} \ln F(X_i) \sim \chi^2_{2n}$$

and

$$-2\sum_{i=1}^{n} \ln[1-F(X_i)] \sim \chi^2_{2n}$$

For this example it is easier to work with the latter, i.e.

$$-2\sum_{i=1}^{n} \ln[1-F(X_i)] \sim \chi^2_{2n}$$

$$-2\sum_{i=1}^{n} \ln[1-F(x_i)] = -2\sum_{i=1}^{n} \ln e^{-\lambda x_i}$$

$$= 2\lambda \sum_{i=1}^{n} x_i = 2n\lambda \bar{x}_n$$

So $\quad 2n\lambda\bar{x}_n \sim \chi^2_{(2n)}$

Using the χ^2 table (Appendix 3, p 850-851)

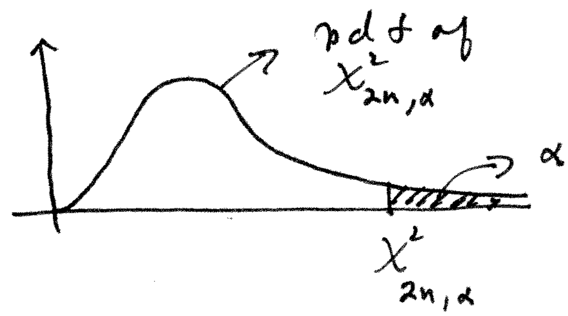
we can find $\chi^2_{(2n),\,0.025}$ and $\chi^2_{(2n),\,0.975}$ such that

$$P\left(\chi^2_{(2n),\,0.975} < 2n\lambda\bar{x}_n < \chi^2_{(2n),\,0.025}\right) = 0.95$$

Thus $\quad \left(\dfrac{\chi^2_{(2n),\,0.975}}{2n\bar{x}_n},\ \dfrac{\chi^2_{(2n),\,0.025}}{2n\bar{x}_n} \right)$

provides a 95% C.I. for λ. Note that

$\chi^2_{(2n),\,\alpha}$ is such that $P\left(\chi^2_{(2n)} > \chi^2_{(2n),\alpha}\right) = \alpha$



- Sample Size Determination —