STAT 5430

Lec 38, W, Apr 30

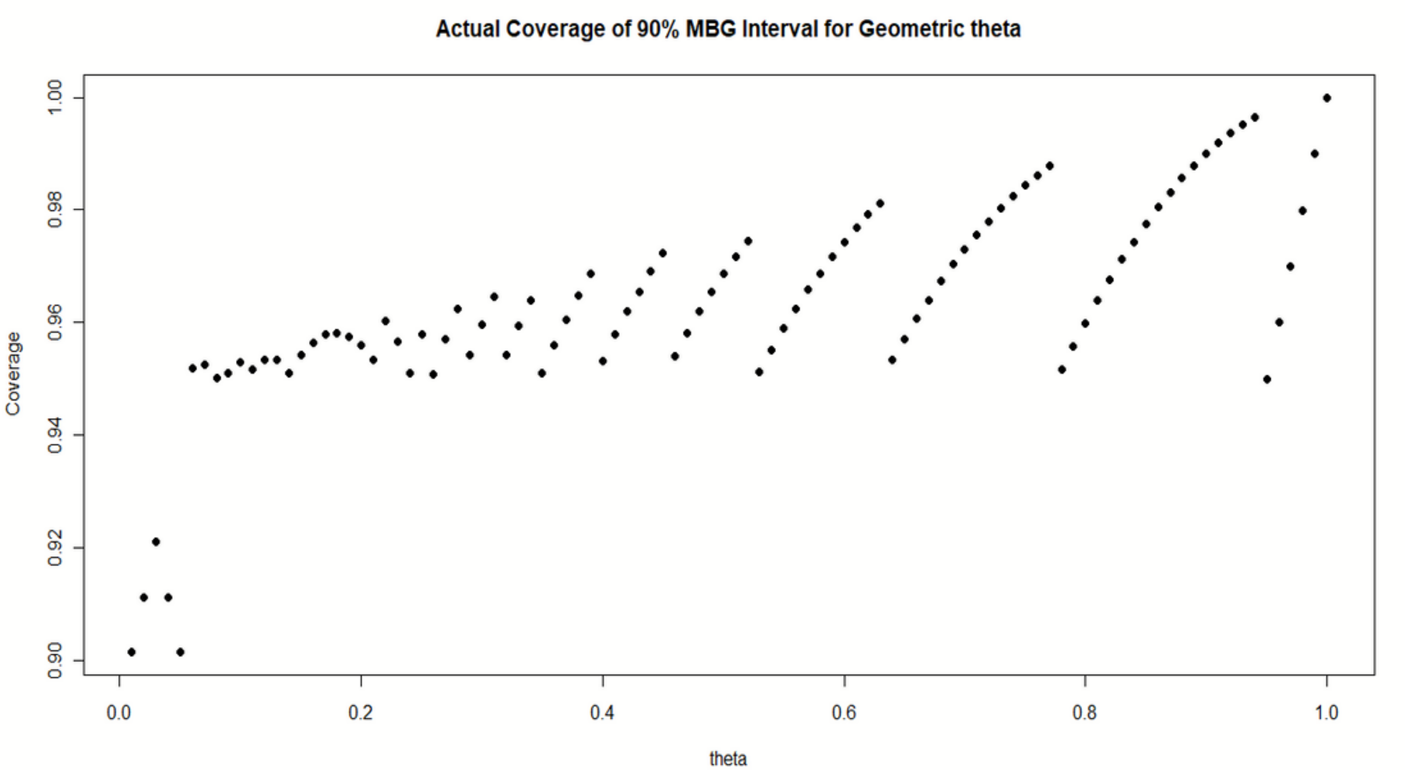
- Homework 9 is assigned & due Sunday, May 4
  but you can submit this on Monday, May 5
- Exam 2 solutions & grading key posted
- Final Exam on Tuesday, May 13, 7:30-9:30 PM
  - Comprehensive - but focus on material since Exam 2 (interval estimation)
  - Formula sheet for new material/interval & 2 formula sheets previous material
    (3 sheets (front/back) each total)
  - Practice Exams

see Canvas

check $P_\theta(\theta \in [\theta_L(T), \theta_U])$ for $\theta \in (0,1)$ numerically

```
theta<-0.01
M<-10000000
CC<-.90


T<-rgeom(M,theta)+1

alpha.2<-(1-CC)/2

LOWER<-1-(1-alpha.2)^{1/T}
UPPER<-1-(alpha.2)^{1/(T-1)}

COVER<-rep(1,M)
COVER[LOWER>theta]<-0
COVER[UPPER<theta]<-0
mean(COVER)
```

```
M<-10000000
CC<-.90

t<-seq(.01,1,.01)
t1<-t*0
n<-length(t)

for(i in 1:n){
theta<-t[i]
 print(i)
T<-rgeom(M,theta)+1

alpha.2<-(1-CC)/2

LOWER<-1-(1-alpha.2)^{1/T}
UPPER<-1-(alpha.2)^{1/(T-1)}

COVER<-rep(1,M)
COVER[LOWER>theta]<-0
COVER[UPPER<theta]<-0

t1[i]<-mean(COVER)
}


plot(t,t1,xlab="theta",ylab="Coverage",main="Actual Coverage of 90% MBG Interval for Geometric Parameter theta",pch=20)
```



Interval Estimation II

Bayes Intervals

Definition: Let $\underset{\sim}{X} = (X_1, \ldots, X_n)$ have joint pdf/pmf $f(\underset{\sim}{x}|\underset{\sim}{\theta})$, $\underset{\sim}{\theta} \in \Theta \subset \Re^p$ and let $\pi(\underset{\sim}{\theta})$ be a prior pdf for $\underset{\sim}{\theta}$ on Θ. Let $f_{\underset{\sim}{\theta}|\underset{\sim}{x}}(\underset{\sim}{\theta})$ be the posterior pdf of $\underset{\sim}{\theta}$ given $\underset{\sim}{X} = \underset{\sim}{x}$. Then, a set $C_{\underset{\sim}{x}}$ is called a $(1-\alpha)$ **credible set** for $\underset{\sim}{\theta}$ if

posterior prob

$$P(\underset{\sim}{\theta} \in C_{\underset{\sim}{x}} | \underset{\sim}{X} = \underset{\sim}{x}) = 1 - \alpha, \ \forall \underset{\sim}{x}$$

credible set $C_{\underset{\sim}{x}} \subset \Theta$ is a guess for where $\underset{\sim}{\theta}$ might be.

or equivalently, that $\underset{\sim}{\theta} \in C_{\underset{\sim}{x}}$

$$\int_{C_{\underset{\sim}{x}}} f_{\underset{\sim}{\theta}|\underset{\sim}{x}}(\underset{\sim}{\theta}) d\underset{\sim}{\theta} = 1 - \alpha, \quad \forall \underset{\sim}{x}.$$

a credible set $C_{\underset{\sim}{x}}$ is NOT confidence region

Not true that $\min_{\underset{\sim}{\theta} \in \Theta} P_{\underset{\sim}{\theta}}(\underset{\sim}{\theta} \in C_{\underset{\sim}{x}}) = 1-\alpha$

Example: Let $X_1, \ldots, X_n$ be iid Poisson(λ), $\lambda > 0$. Let the prior for λ on $(0, \infty)$ be Exponential(1). $\leftarrow$ gamma(1,1)

Then, the posterior dist. of λ given $\underset{\sim}{X}$ is

$\text{gamma}(1+\sum_{i=1}^n x_i, (n+1)^{-1})$ [check]

Hence, a $(1-\alpha)$ credible set/interval for λ is given by $[L, U]$ where

$P(\lambda \in [L, U] | \underset{\sim}{X}) = 1-\alpha$ (posterior prob)

Note: $\lambda | \underset{\sim}{X} \sim \text{gamma}(1+\sum_{i=1}^n x_i, (n+1)^{-1})$

$\Rightarrow \dfrac{2\lambda}{(n+1)^{-1}} \Big| \underset{\sim}{X} \sim \chi^2_{2(1+\sum_{i=1}^n x_i)}$ (posterior dist)

So, $1-\alpha = P(L \le \lambda \le U \mid \underset{\sim}{X})$

$$= P\Big(\underbrace{\frac{2L}{(n+1)^{-1}}}_{\chi^2_{2(1+\Sigma x_i);\frac{\alpha}{2}}} \le \underbrace{\frac{2\lambda}{(n+1)^{-1}}}_{\chi^2_{2(1+\Sigma x_i)}} \le \underbrace{\frac{2U}{(n+1)^{-1}}}_{\chi^2_{2(1+\Sigma x_i);1-\frac{\alpha}{2}}} \,\Big|\, \underset{\sim}{X}\Big)$$

So, finally, $[L, U]$, where

$$L = \frac{\chi^2_{2(\Sigma x_i+1);\frac{\alpha}{2}}}{2(n+1)} \quad \& \quad U = \frac{\chi^2_{2(\Sigma x_i+1);1-\frac{\alpha}{2}}}{2(n+1)},$$

is a $(1-\alpha)$ credible interval for λ.

Interval Estimation II

Bayes Intervals

Definition: A **highest posterior density** (HPD) credible set of level $(1-\alpha)$ is a set of the form

← pick "cut-off" C

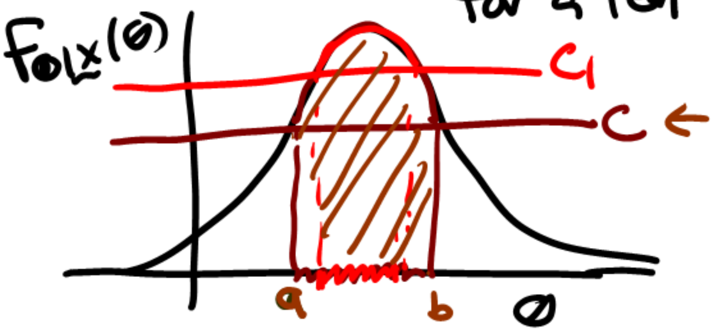
$$C_{\underset{\sim}{x}} = \{\underset{\sim}{\theta} \in \Theta : f_{\underset{\sim}{\theta}|\underset{\sim}{x}}(\underset{\sim}{\theta}) \geq C\}, \text{ for } \textbf{some } C > 0$$

such that $\underbrace{P(\underset{\sim}{\theta} \in C_{\underset{\sim}{x}}|\underset{\sim}{X} = \underset{\sim}{x})}_{\text{posterior prob}} = 1 - \alpha, \ \forall \underset{\sim}{x}$.

has "right" post. prob to be a $(1-\alpha)$ credibility set

Discussion: Why do this?

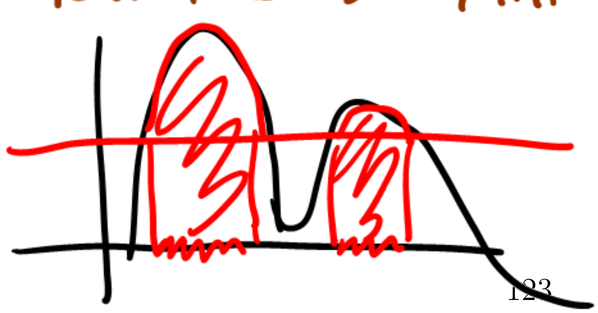
Consider posterior density $f_{\theta|\underset{\sim}{x}}(\theta)$ for a real-valued θ, which is unimodal



pick c so that $\int_a^b f_{\theta|\underset{\sim}{x}}(\theta)\,d\theta = 1-\alpha$ ⇒ gives the HPD credible set as $C_{\underset{\sim}{x}} = [a,b]$

Note: Where $f_{\theta|\underset{\sim}{x}}(\theta)$ is "high", want to "pack in" an area of $1-\alpha$ over a short region of θ.

So, HPD credible sets achieve $(1-\alpha)$ posterior coverage but tend to be small/informative sets for θ (guesses for θ)



Interval Estimation II

Bayes HPD Intervals: Illustration

Example: Let $X_1, \ldots, X_n$ be iid $N(\theta, \sigma^2)$ with $\theta \in \mathbb{R}$ and known $\sigma^2 > 0$. Suppose a prior distribution for θ is $N(\mu, \tau^2)$ for some known $\mu \in \mathbb{R}, \tau^2 > 0$.

$\Rightarrow$ posterior distribution of θ given $\underset{\sim}{X}$ is

$$\theta \mid \underset{\sim}{X} \sim \text{Normal}\left(\mu_{\theta \mid \underset{\sim}{X}}, \sigma^2_{\theta \mid \underset{\sim}{X}}\right)$$

unimodal

depend on $\underset{\sim}{X}, n, \mu, \tau^2$
(done this before)

Find a $(1-\alpha)$ HPD credible set for θ:

$$C_{\underset{\sim}{X}} = \{\theta : f_{\theta \mid \underset{\sim}{X}}(\theta) \geq c\} = \left\{\theta : \frac{1}{\sqrt{2\pi}\,\sigma_{\theta \mid \underset{\sim}{X}}} e^{-\frac{(\theta - \mu_{\theta \mid x})^2}{2\sigma^2_{\theta \mid \underset{\sim}{X}}}} \geq c\right\}$$

$$= \left\{\theta : \left|\frac{\theta - \mu_{\theta \mid \underset{\sim}{X}}}{\sigma_{\theta \mid \underset{\sim}{X}}}\right| \leq c_1\right\}$$