(A) Assume "δ", a margin of equivalence to modify hypothesis, ~~i.e. a margi~~

$H_0: \theta \in [\theta_0 \pm \delta]$, $H_a: \theta \notin [\theta_0 \pm \delta]$, $P_{val} = P(H_0|X) = P(\theta \in [\theta_0 - \delta, \theta_0 + \delta \,|\, X])$ ← fair coin example

This modifies our hypothesis to include "close" values e.g. $\theta_0 = 0.5$, $\delta = 0.01$ $\theta \in [0.49, 0.51]$

for $n=100$, $x=43$, $\alpha = 5\%$. $P(\theta) \sim U(0,1) \Rightarrow P(\theta|x) = Beta(44,58)$

$H_0: \theta \in [0.5 \pm 0.01]$, $H_A: \theta \notin [0.5 \pm 0.01]$

$P_{val} = P(H_0|X) = pbeta(0.51, 44, 58) - pbeta(0.49, 44, 58) = 0.06$

F

Downside! new decision parameter introduced (which btw is arbitrary)

(B) Constraint $CR_{\theta,1-\alpha}$

if $\theta_0 \in CR_{\theta,1-\alpha} \Rightarrow$ Retain, else reject

eg. $CR_{\theta,95\%} = [qbeta(.025, 44, 58), qbeta(.975, 44, 58)] = [.337, .528] \Rightarrow$ Retain H_0

Downside! No Baysian $P_{val} = P(H_0|X)$

GOALS OF STATISTICS

(I) Inference
- (a) Point estimate
- (b) Confidence Set
- (c) Theory Testing

(II) Predict

Declare χ ← beyond scope of class

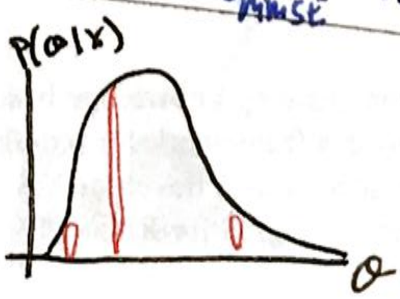
- We observe $X_1 \dots X_n$ (data)
- We want to predict X^* given our data
- Obviously you cant KNOW X^*, so we make a model
- $X \sim P(X|\theta)$ But we dont KNOW θ, we infer it!
- So, the next best thing is using $X_1 \dots X_n$ to infer θ for $X^*|\vec{X}_n$.
- The naive thing to do is $X^*|X \sim p(X|\hat{\theta}_{MLE})$ which for $n=100$, $x=43$
  $X^*|X \sim Bern(0.43)$ Naive i.e. simplistic
  This is silly b/c $n=5$, $x=0$ gives $\hat{\theta}_{MLE} = 0$ which says ALWAYS get failures...
- The Bayesian Way: $P(\theta) = Beta(1,1)$ which will update $X^*|\vec{X} \sim P(X|\theta = \hat{\theta}_{MMSE})$
  $n=5, x=0 \rightarrow X^*|\vec{X} \sim Bern(1/7)$

But why use the posterior mean $\hat{\theta}_{MMSE}$ as opposed to $\hat{\theta}_{MAP}$, $\hat{\theta}_{MMAE}$...

$$P(X^*|X) = \int_\Theta P(X^*, \theta|X)\, d\theta$$

def of conditional probability

$$= \int_\Theta \underbrace{P(X^*|\theta, X)}_{p(X^*|\theta) \text{ for a specific } \theta} P(\theta|X)\, d\theta = \overbrace{\int_\Theta \underbrace{P(X^*|\theta)}_{\text{likelihood}} \underbrace{P(\theta|X)}_{\text{posterior, everything we know about } \theta} d\theta}^{\text{wieghted mean}}$$



[POSTERIOR PREDICTIVE DISTRIBUTION]

We get a possible θ from the Posterior to use in the likelihood, and then averaging each possibility by how likely we think it is, getting the above distribution. Each likelihood is wieghed by its relative θ.

Discrete version of Posterior Predictive Distribution

$\sum_{\theta \in \Theta} P(x^*|\theta)\, P(\theta|X)$ and use it for Binomial

For $\Theta = \{0.5, 0.75\}$ $X = \langle 0,1,1 \rangle$ $\longrightarrow$

$$P(x^*|X) = \underbrace{P(x^*|\theta=0.5)}_{(0.5)^{x^*}(0.5)^{1-x^*} \,=\, 0.5}\ \underbrace{P(\theta=0.5\,|\,\langle 0,1,1\rangle)}_{0.47\,*}$$
$$+\ \underbrace{P(x^*|\theta=0.75)}_{(0.75)^{x^*}(0.25)^{1-x^*}}\ \underbrace{P(\theta=0.75\,|\,\langle 0,1,1\rangle)}_{0.53\,*}$$

* from past notes

use PMF Bern

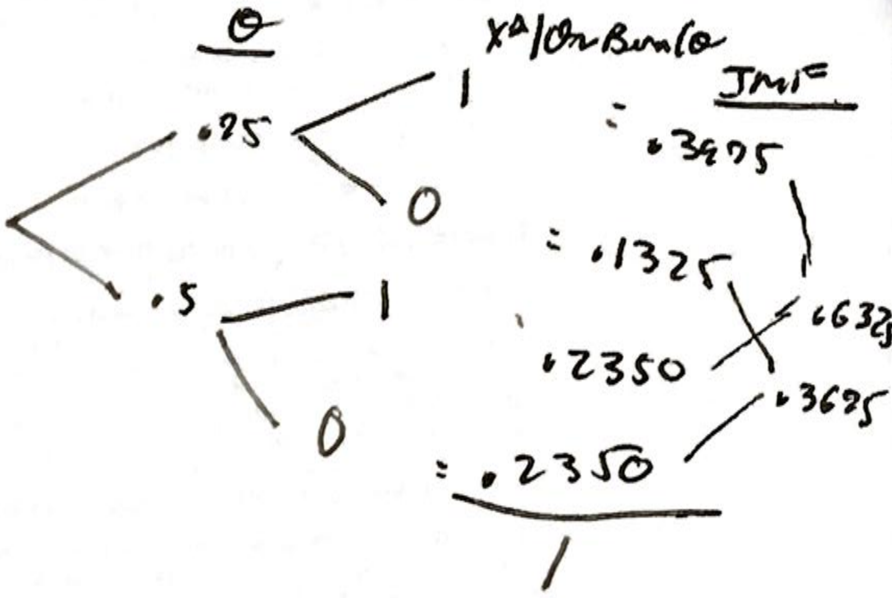
$$P(x^*|X) = (0.5)(0.47) + (0.75)^{x^*}(0.25)^{1-x^*}(0.53)$$

This doesn't look special BUT!

We know $P(x^*|X) \sim$ Bernoulli (b/c $Supp[x^*] = \{0,1\}$)

this means $P(x^*=1\,|X) = (0.5)(0.47) + (0.75)^1(0.25)^0(0.53) = .6325$

which means $P(x^*|X) \sim$ Bernoulli$(.6325)$



This is easy to see discrete... How about for continuous Θ?

$F \sim$ Binomial, $P(\theta) = Beta(\alpha, \beta)$

$$P(x^*|X) = \int_\Theta P(x^*|\theta)P(\theta|X)\,d\theta = \int_0^1 \theta^{x^*}(1-\theta)^{1-x^*}\left(\frac{1}{B(\alpha+x,\, n-x+\beta)}\,\theta^{x+\alpha-1}(1-\theta)^{n-x+\beta-1}\right)d\theta$$

margin out θ next

$$\frac{1}{B(x+\alpha,\, n-x+\beta)} \int_0^1 \theta^{(x^*+x+\alpha)-1}(1-\theta)^{(n-x+\beta-x^*+1)-1}\,d\theta = B$$

$$\hookrightarrow = \boxed{\frac{B(x^*+x+\alpha,\, n-x+\beta-x^*+1)}{B(x+\alpha,\, n-x+\beta)}}$$

We know $P(x^*|X) \sim$ Bern( ) so lets test $x^*=1$

$$P(x^*=1|X) = \frac{B(x+\alpha+1,\, n-x+\beta)}{B(x+\alpha,\, n-x+\beta)} = \frac{\frac{(x+\alpha)}{(n+\alpha+\beta)}\overbrace{\frac{\Gamma(x+\alpha)\Gamma(n-x+\beta)}{\Gamma(n+\alpha+\beta)}}^{B(x+\alpha,\, n-x+\beta)}}{B(x+\alpha,\, \cancel{n-x+\beta})} = \boxed{\frac{x+\alpha}{(n+\alpha+\beta)}} = \hat{\theta}_{MMSE}$$

$$\therefore P(x^*|X) \sim Bern\left(\frac{x+\alpha}{(n+\alpha+\beta)}\right)$$

another reason why this is the preferred Bayesian Estimator

$\boxed{\text{Conclude Midterm 1 material}}$