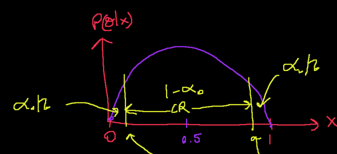
$x=1,\ n=2,\ \mathcal{F}: Bin(n,\theta)$

$P(\theta) = Beta(1,1) \Rightarrow P(\theta|x) = Beta(2,2)$

Produce a 95% credible region for θ.



$$CR_{\theta,1-\alpha_0} := \left[Q\left[\theta|x, \frac{\alpha_0}{2}\right], Q\left[\theta|x, 1-\frac{\alpha_0}{2}\right] \right]$$

$$= \left[qbeta(2.5\%, 2, 2),\ qbeta(97.5\%, 2, 2)\right]$$

$$= [0.094,\ 0.906]$$

$$P(\theta \in [.094, .906] \mid x) = 95\%$$

This is a real probability statement! The CI approach cannot give you such a statement! The CR is highly interpretable. Also, the CR is a proper subset of Θ for alpha_0 > 0. This is not always true with CI's e.g. here's the 95% CI for this data:

$$CI_{\theta,95\%} = \left[0.5 \pm 1.96\sqrt{\frac{0.5 \cdot 0.5}{2}}\right] = [-0.21,\ 1.21] \not\subset \Theta = (0,1)$$

The above CR is technically a two-sided CR. You can also create one-sided (i.e. left-sided or right-sided) CR's:

$CR_{L,\theta,1-\alpha_0} :=$ [smallest value in Θ or -oo, $Q[\theta|x, 1-\alpha_0]$]

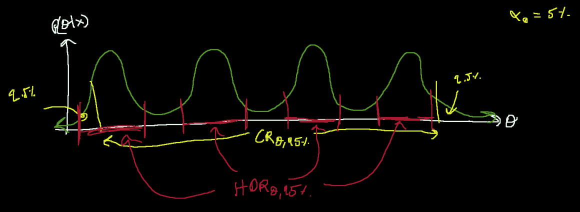
e.g. in our dataset

$$= [0,\ \underset{qbeta(.95,2,2)}{0.865}] \Rightarrow P(\theta < 0.865 \mid x) = 95\%$$

$CR_{R,\theta,1-\alpha_0} :=$ [$Q[\theta|x, \alpha_0]$, largest value in Θ or +oo]

$$= [qbeta(.05, 2, 2),\ 1] = [.136,\ 1] \Rightarrow P(\theta \geq .136 \mid x) = 95\%$$

Another approach (which we will see but not study further) is called the high density region (HDR) approach. Consider the following posterior for θ:



$P(\theta \in HDR_{\theta,95\%}) = 95\%$ but it has *minimum* width (the sum of the interval pieces is minimum).

Sometimes the CR = HDR (e.g. in unimodal posteriors).

Disadvantages of the HDR approach. (1) it can be non-contiguous i.e. in pieces! (2) it's computationally intense (3) no L or R intervals.

Bayesian hypothesis testing. We can immediately compute the following quantities:

Bayes-ian p-value $:= P(H_0|x),\quad P(H_a|x)$

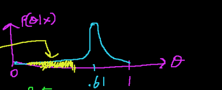
threshold of "sufficient evidence"

$$\text{if } P(H_0|x) < \alpha_0 \Rightarrow \text{Reject } H_0.$$

Let's recreate the hypothesis testing example from Lec 3. n = 100 flips of a coin where x = 61 were heads. Test if the coin is unfairly weighted towards heads at a 5% significance level.

$H_a: \theta > 0.5 \Rightarrow H_0: \theta \leq 0.5.$ Assume $P(\theta) = Beta(1,1)$

$$P(\theta|x) = Beta(62, 40)$$



$$P(H_0|x) = P(\theta \leq 0.5|x) = \int_0^{0.5} P(\theta|x)\,d\theta = \int_0^{0.5} \frac{1}{B(62,40)} \theta^{61}(1-\theta)^{31}\,d\theta$$

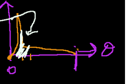
$$= pbeta(0.5, 62, 40) = .014 = 1.4\% \Rightarrow \text{Reject } H_0 \text{ i.e.}$$

coin is unfairly weighted towards heads.

$n=200 \qquad x=37 \qquad \mathcal{F}: Bin(n, \theta)$

Uber driver does 200 rides and gets 37 non-5-star ratings. If his true proportion of non-5-star ratings is more than 25%, then Uber policy is to not fire the driver. Prove he should be fired (or not) at a 5% significance level.

$H_a: \theta > 25\% \Rightarrow H_0: \theta \leq 25\%.\quad P(\theta) = Beta(1,1)$



$$\Rightarrow P(\theta|x) = Beta(38, 164)$$

$$P(H_0|x) = P(\theta \leq .25|x) = \int_0^{0.25} \ d\theta$$

$$= pbeta(.25, 38, 164) = .983 > 5\% \Rightarrow \text{Retain } H_0. \text{ Don't fire him.}$$

$n=100 \qquad x=43$

Let's test the coin again. Flip 100 times and get 43 heads. Test if the coin is in unfair at 5% significance. $H_a: \theta \neq 0.5 \Rightarrow H_0: \theta = 0.5$

$$P(\theta) = Beta(1,1) \Rightarrow P(\theta|x) = Beta(44, 58)$$

$$P(H_0|x) = P(\theta = 0.5|x) = 0 \Rightarrow \text{Reject } H_0 \text{ always ?? YES...}$$

Using this approach, two-sided tests are always rejected (if the posterior is continuous). Does this make sense? Any infinitely precise theory of theta is wrong in the real world. A coin is never exactly 50.000000000....% likely to flip heads.

So we need to slightly reframe our hypotheses using the notion of "margin of equivalence" called δ (delta).

$$H_a: \theta \notin [\theta_0 \pm \delta] \Rightarrow H_0: \theta \in [\theta_0 \pm \delta]$$