Recall the example $\underset{\text{Data}}{X} \sim \text{Bin}(n, \underset{\text{Parameter}}{\theta})$

Want to incorperate knowledge that $\theta \in \{0.25, 0.5, 0.75\}$
with prior distribution

	$\mathbb{P}(\theta = 0.25)$	$\mathbb{P}(\theta = 0.5)$	$\mathbb{P}(\theta = 0.75)$	
Conductor:	p	$10p$	p	Informative prior
friend	p	p	p	Noninformative prior

Used Bayes' Rule to attain the posterior

$$\mathbb{P}(\theta \mid X) = \frac{\overbrace{\mathbb{P}(X \mid \theta)}^{\text{likelihood}} \overbrace{\mathbb{P}(\theta)}^{\text{prior}}}{\underbrace{\mathbb{P}(X)}_{\text{evidence}}}$$

Under the noninformative prior

$$\mathbb{P}(\theta = .5 \mid X) = \frac{\mathbb{P}(X \mid \theta = .5)\,\mathbb{P}(\theta = .5)}{\sum_{t \in \Theta} \mathbb{P}(X \mid \theta = t)} \simeq 0.05$$

$$\mathbb{P}(\theta = .5 \mid X) = \text{———}''\text{———} \simeq 0.95$$

Under the informative prior you get much different results.

What is we know observe an extra coin flip? Say $Y = 1$.

How would we update our beliefs?

$$\mathbb{P}(\theta \mid X,Y) = \frac{\mathbb{P}(X,Y \mid \theta)\,\mathbb{P}(\theta)}{\mathbb{P}(X,Y)} = \frac{\mathbb{P}(Y \mid \theta, X)}{\mathbb{P}(Y \mid X)} \underbrace{\frac{\mathbb{P}(X \mid \theta)\,\mathbb{P}(\theta)}{\mathbb{P}(X)}}_{\text{original posterior}}$$

$$= \frac{\mathbb{P}(Y \mid \theta, X)\,\mathbb{P}(\theta \mid X)}{\mathbb{P}(Y \mid X)}$$

$$\propto \mathbb{P}(Y \mid \theta)\,\mathbb{P}(\theta \mid X)$$

Therefore the posterior now plays the role of the prior distribution.

This is known as the sequential update.

Ex: $\mathbb{P}(\theta = .5 \mid X, Y) = \dfrac{\mathbb{P}(Y=1 \mid \theta = .5)\,\mathbb{P}(\theta = .5 \mid X)}{\sum_{t \in \Theta} \mathbb{P}(Y=1 \mid \theta = t)\,\mathbb{P}(\theta = t \mid X)} = \dfrac{(0.5)(0.05)}{\sum_{t \in \Theta} t\,\mathbb{P}(\theta = t \mid X)}$

Calculating the odds ratio

$$\frac{\mathbb{P}(\theta = .75 \mid X, Y)}{\mathbb{P}(\theta = .5 \mid X, Y)} = \frac{.75\,\mathbb{P}(\theta = .75 \mid X)}{.5\,\mathbb{P}(\theta = .5 \mid X)} \approx \left(\frac{3}{2}\right) \cdot (.96) \approx 1.44$$

Using this posterior for predictive inference of a new coin flip $\tilde{X}$.

$$\mathbb{P}(\tilde{X} \mid X) = \sum_{\theta} \mathbb{P}(\tilde{X}, \theta \mid X) = \sum_{\theta} \mathbb{P}(\tilde{X} \mid \theta, X)\,\mathbb{P}(\theta \mid X) = \sum_{\theta} \mathbb{P}(\tilde{X} \mid \theta)\,\mathbb{P}(\theta \mid X)$$

Posterior predictive distribution (probability)

Ex: $\sum_{t \in \Theta} \mathbb{P}(\tilde{X} \mid \theta = t)\,\mathbb{P}(\theta = t \mid X) = \sum_{t \in \Theta} t\,\mathbb{P}(\theta = t \mid X) = \mathbb{E}_{\theta \mid X}[\theta]$

Ex: Suppose X_1, X_2 iid with $\mathbb{P}(X_i = \theta - 1) = \mathbb{P}(X_i = \theta + 1) = \frac{1}{2}$

We observe the sample $X = (X_1, X_2)$ and choose the estimator

$$\hat{\theta}(X) = \begin{cases} \frac{1}{2}(X_1 + X_2) & X_1 \neq X_2 \\ X_1 - 1 & X_1 = X_2 \end{cases}$$

x_1	x_2	Prob	
$\theta-1$	$\theta-1$	$1/4$	✗
$\theta-1$	$\theta+1$	$1/4$	✓
$\theta+1$	$\theta-1$	$1/4$	✓
$\theta+1$	$\theta+1$	$1/4$	✓

$\mathbb{P}(\hat{\theta}(X) = \theta) = \frac{3}{4}$

So by definition $\hat{\theta}(X)$ is a 75% confidence interval for θ

But if we know the data is different, then we should be 100% confident.

and similarly if they are the same, we should be 50% confident.

Rmk: Those estimators are <u>conditional on the data.</u>

Ex: $X \in \{1, 2, 3\}$

	1	2	3
$\mathbb{P}(X \mid \theta = 0)$	.005	.005	.99
$\mathbb{P}(X \mid \theta = 1)$	.005	.985	.01

Test $H_0: \theta = 0$ $H_A: \theta = 1$ take $R = \{1, 2\}$

$\alpha = \mathbb{P}_0(X \in R) = 0.01$

$\text{power} = 1 - \beta = \mathbb{P}_A(X \in R) = 0.99$

Suppose you observe $X = 1$. Then we are confident that H_0 is false

The LR $\frac{P(X=1|\theta=0)}{P(X=1|\theta=1)} = 1$ which hints there is no support for H_0.

Should probably be a toss-up.

The *likelihood principle*: Making inferences about parameters after observing data all relevant experimental information is found in the likelihood function.

Ex: Under the coin tossing example consider $H_0: \theta = 1/2$. $H_a: \theta > 1/2$

Suppose $n = 12$ and 9 are correct.

Could have observed $X \sim Bin(n=12, \theta)$, observe $X = 9$ $(*)$

Also could have come from a negative binomial $X \sim NegBin(r=3, \theta)$ $(**)$

$$P_*(X|\theta) = \binom{n}{x} \theta^x (1-\theta)^{n-x} = \binom{12}{9} \theta^9 (1-\theta)^3$$

$$P_{**}(X|\theta) = \binom{r+x-1}{x} \theta^x (1-\theta)^r = \binom{11}{9} \theta^9 (1-\theta)^3$$

The nominal rejection region is $R = \{X \geq 9\}$

$\left.\begin{array}{l} pval = P_*(X \geq 9 | \theta = 1/2) \approx 0.073 \\ pval = P_{**}(X \geq 9 | \theta = 1/2) \approx 0.033 \end{array}\right\}$ Different conclusion based on $\alpha = .05$ rejection region.

Rmk: As $P(\theta|X) \propto \underbrace{P(X|\theta)}_{\text{experimental info — still follows the likelihood principle}} P(\theta)$