18.05 Lecture 23
April 6, 2005

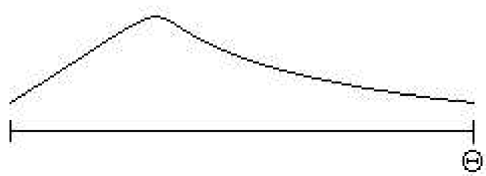
Estimation Theory:
If only 2 outcomes: Bernoulli distribution describes your experiment.
If calculating wrong numbers: Poisson distribution describes experiment.
May know the type of distribution, but not the parameters involved.

A sample (i.i.d.) $X_1, ..., X_n$ has distribution $\mathbb{P}$ from the family of distributions:
$\{\mathbb{P}_\theta : \theta \in \Theta\}$
$\mathbb{P} = \mathbb{P}_{\theta_0}, \theta_0$ is unknown
Estimation Theory - take data and estimate the parameter.
It is often obvious based on the relation to the problem itself.

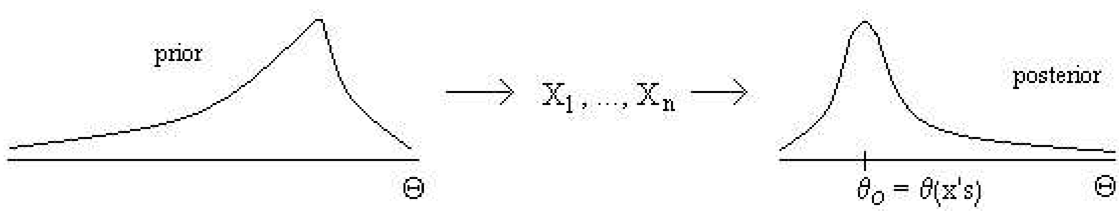
Example: B(p), sample: 0 0 1 1 0 1 0 1 1 1
$p = \mathbb{E}(X) \leftarrow \overline{x} = 6/10 = 0.6$

Example: $E(\alpha), \alpha e^{-\alpha x}, x \geq 0, \mathbb{E}(X) = 1/\alpha$.
Once again, parameter is connected to the expected value.
$1/\alpha = \mathbb{E}(X) \leftarrow \overline{x}, \alpha \approx 1/\overline{x}$ - estimate of alpha.

Bayes Estimators: - used when intuitive model can be used in describing the data.



$X_1, ..., X_n \sim \mathbb{P}_{\theta_0}, \theta_0 \in \Theta$
Prior Distribution - describes the distribution of the set of parameters (NOT the data)
$f(\theta)$ - p.f. or p.d.f. $\rightarrow$ corresponds to intuition.
$\mathbb{P}_0$ has p.f. or p.d.f.; $f(x|\theta)$
Given $x_1, ..., x_n$ joint p.f. or p.d.f.: $f(x_1, ..., x_n|\theta) = f(x_1|\theta) \times ... \times f(x_n|\theta)$
To find the Posterior Distribution - distribution of the parameter given your collected data.
Use Bayes formula:



$$f(\theta|x_1, ..., x_n) = \frac{f(x_1, .., x_n|\theta)f(\theta)}{\int f(x_1, ..., x_n|\theta)f(\theta)d\theta}$$

The posterior distribution adjusts your assumption (prior distribution) based upon your sample data.

Example: $B(p), f(x|p) = p^x(1-p)^{1-x}$;

$$f(x_1, ..., x_n|p) = \Pi p^{x_i}(1-p)^{1-x_i} = p^{\sum x_i}(1-p)^{n-\sum x_i}$$

Your only possibilities are p = 0.4, p = 0.6, and you make a prior distribution based on the probability that the parameter p is equal to each of those values.
Prior assumption: f(0.4) = 0.7, f(0.6) = 0.3
You test the data, and find that there are are 9 successes out of 10, $\hat{p} = 0.9$
Based on the data that give $\hat{p} = 0.9$, find the probability that the actual p is equal to 0.4 or 0.6.
You would expect it to shift to be more likely to be the larger value.
Joint p.f. for each value:

$$f(x_1, ..., x_{10}|0.4) = 0.4^9(0.6)^1$$

$$f(x_1, ..., x_{10}|0.6) = 0.6^9(0.4)^1$$

Then, find the posterior distributions:

$$f(0.4|x_1, ..., x_n) = \frac{(0.4^9(0.6)^1)(0.7)}{(0.4^9(0.6)^1)(0.7) + (0.6^9(0.4)^1)(0.3)} = 0.08$$

$$f(0.6|x_1, ..., x_n) = \frac{(0.6^9(0.4)^1)(0.3)}{(0.4^9(0.6)^1)(0.7) + (0.6^9(0.4)^1)(0.3)} = 0.92$$

Note that it becomes much more likely that p = 0.6 than p = 0.4

Example: B(p), prior distribution on [0, 1]
Choose any prior to fit intuition, but simplify by choosing the conjugate prior.

$$f(p|x_1, ..., x_n) = \frac{p^{\Sigma x_i}(1-p)^{n-\Sigma x_i} f(p)}{\int (...)dp}$$

Choose f(p) to simplify the integral. Beta distribution works for Bernoulli distributions.
Prior is therefore:

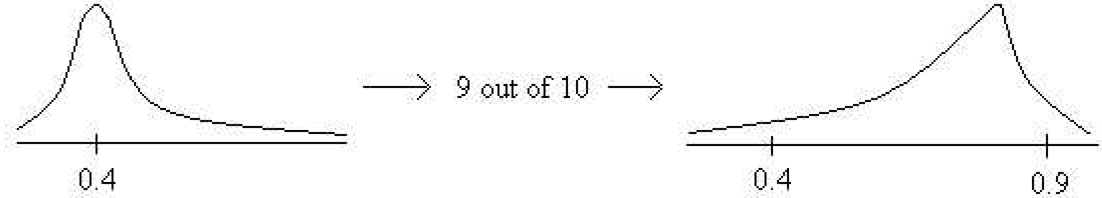
$$f(p) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1}, 0 \le p \le 1$$

Then, choose α and β to fit intuition: makes $\mathbb{E}(X)$ and $\mathrm{Var}(X)$ fit intuition.

$$f(p|x_1...x_n) = \frac{\Gamma(\alpha + \sum x_i + \beta + n - \sum x_i)}{\Gamma(\alpha + \sum x_i)\Gamma(\beta + n - \sum x_i)} \times p^{(\alpha+\sum x_1)-1}(1-p)^{(\beta+n-\sum x_i)-1}$$

Posterior Distribution = Beta$(\alpha + \sum x_i, \beta + n - \sum x_i)$
The conjugate prior gives the same distribution as the data.

Example:



$B(\alpha, \beta)$ such that $\mathbb{E}X = 0.4, \text{Var}(X) = 0.1$
Use knowledge of parameter relations to expectation and variance to solve:

$$\mathbb{E}X = 0.4 = \frac{\alpha}{\alpha + \beta}, \text{Var}(X) = 0.1 = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}$$

The posterior distribution is therefore:

$$\text{Beta}(\alpha + 9, \beta + 1)$$

And the new expected value is shifted:

$$\mathbb{E}X = \frac{\alpha + 9}{\alpha + \beta + 10}$$

Once this posterior is calculated, choose the parameters by finding the expected value.

Definition of Bayes Estimator:
Bayes estimator of unknown parameter θ_0 is $\theta(X_1, ..., X_n)$ = expectation of the posterior distribution.

Example: B(p), prior $\text{Beta}(\alpha, \beta), X_1, ..., X_n \rightarrow$ posterior $\text{Beta}(\alpha + \sum x_i, \beta + n - \sum x_i)$

$$\text{Bayes Estimator: } \frac{\alpha + \sum x_i}{\alpha + \sum x_i + \beta + n - \sum x_i} = \frac{\alpha + \sum x_i}{\alpha + \beta + n}$$

To see the relation to the prior, divide by n:

$$= \frac{\alpha/n + \overline{x}}{\alpha/n + \beta/n + 1}$$

Note that it erases the intuition for large n.
The Bayes Estimator becomes the average for large n.

** End of Lecture 23