



# PRIOR DISTRIBUTIONS

# Leverage



# i has it

# WHY USING INFORMED PRIORS?

- ◉ In a Bayesian analysis we can use informed priors or uninformed priors
  - Uninformed priors do not add any information to the collected data
  - Informed priors add to the collected data the information that possibly already exists (due to previous experiments or due to generally available knowledge)
- ◉ The use of uninformed priors (i.e., to rely only on the collected data) is a way to prevent the inclusion of biases in our analysis
- ◉ Nonetheless, whenever some knowledge already exists it is a waste of resources not to take advantage of it

# WHY USING INFORMED PRIORS?

- ◉ Furthermore, especially with very undetermined systems (like our system of 1 population and 1 test), the use of uninformed priors would lead to very flat posterior distributions of the estimated parameters,
- ◉ even wider than what we could reasonably expect for the performance of any lab test



# GENERATION OF INFORMED PRIORS?

- ⦿ The problem with past knowledge is that
- ⦿ we need the parameters  $\alpha$  and  $\beta$  of a Beta distribution
- ⦿ and published data (or experience) are almost never in the form of a Beta distribution.
- ⦿ So, we need a method to translate our information into the parameters of a Beta distribution





ESTIMATION OF  $\alpha$  AND  $\beta$

# ESTIMATION OF $\alpha$ AND $\beta$

- From the statistics we know that for a Beta distribution:

$$E(x) = \frac{\alpha}{\alpha + \beta}$$

$$Var(x) = \frac{\alpha * \beta}{(\alpha + \beta)^2 * (\alpha + \beta + 1)}$$

- Knowing the Expected value and the Variance of a Beta distribution we can estimate the two parameters  $\alpha$  and  $\beta$

# ESTIMATION OF $\alpha$ AND $\beta$

- Formulas for  $\alpha$  and  $\beta$  are:

$$\alpha = \frac{[Exp(x)]^2 - [Exp(x)]^3 - Exp(x) * Var(x)}{Var(x)}$$

$$\beta = \frac{Exp(x) - 2 * [Exp(x)]^2 + [Exp(x)]^3 - Var(x) + Exp(x) * Var(x)}{Var(x)}$$

# ESTIMATION OF $\alpha$ AND $\beta$

- ⊙ Data that we may have are:
  - From scientific literature, usually an average and a standard error or a standard deviation
  - From the elicitation of expert opinion
    - the most likely value
    - an upper ceiling for the possible values of our parameter or
    - a baseline below which the parameter value is unlikely to be
- ⊙ Programs exist that may assist in the estimation of our parameters of interest, but the simplest way is to use an Excel spreadsheet

# DATA AVAILABLE: MEAN, AND BASELINE OR CEILING



calc to find a prior.xlsx - Microsoft Excel

Home Insert Page Layout Formulas Data Review View Add-Ins

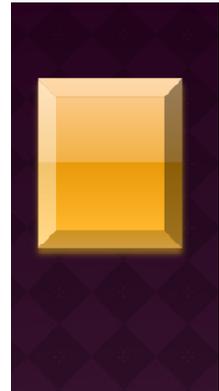
Clipboard Font Alignment Number Styles Cells Editing

	A	B	C	D	E	F	G	H	I	J	K	L
1	average=	0.27										
2	5 or 95 perc=	0.5										
3												
4												
5	$\sigma$ =	=ABS(average-5or95perc)/ABS(NORMINV(0.05,0,1))										
6	Var=	= $\sigma^2$										
7												
8	a=	=(average^2-average^3-average*Var)/Var										
9	b=	=(average-2*average^2+average^3-Var+average*Var)/Var										
10												
11			average= $=\alpha/(\alpha+\beta)$									
12	x	cum p	5 p=	=LOOKUP(0.05,[cum p],[x])								
13		0	95 p=	=LOOKUP(0.95,[cum p],[x])								
14		0.01	=BETADIST(x, $\alpha$ , $\beta$ )									
15		0.02	=BETADIST(x, $\alpha$ , $\beta$ )									
16		0.03	=BETADIST(x, $\alpha$ , $\beta$ )									
17		0.04	=BETADIST(x, $\alpha$ , $\beta$ )									

Sheet1 Sheet1 (2) Sheet2 Sheet3

Ready 147%

# DATA AVAILABLE: MEAN, AND STANDARD DEVIATION



calc to find a prior.xlsx - Microsoft Excel

Home Insert Page Layout Formulas Data Review View Add-Ins

Normal Page Layout Page Break Preview Custom Views Full Screen

Workbook Views

Ruler Formula Bar Gridlines Headings Message Bar

Show/Hide

Zoom 100% Zoom to Selection

Zoom

New Window Arrange All Freeze Panes Hide Split Unhide

View Side by Side Synchronous Scrolling Reset Window Position

Window

Save Workspace Switch Windows

Macros

H2

	A	B	C	D	E	F	G	H	I	J	K	L
1	average=	0.27										
2	2.5 p	0.2										
3	97.5 p	0.35										
4												
5	$\sigma$ =	=(97.5p-2.5p)/(2*1.96)										
6	Var=	= $\sigma^2$										
7												
8	$\alpha$ =	=(average^2-average^3-average*Var)/Var										
9	$\beta$ =	=(average-2*average^2+average^3-Var+average*Var)/Var										
10												
11			average=	= $\alpha/(\alpha+\beta)$								
12	x	cum p	2.5 p=	=LOOKUP(0.025,[cum p],[x])								
13		0	=BETADIS	97.5 p=	=LOOKUP(0.975,[cum p],[x])							
14		0.01	=BETADIST(x, $\alpha$ , $\beta$ )									
15		0.02	=BETADIST(x, $\alpha$ , $\beta$ )									
16		0.03	=BETADIST(x, $\alpha$ , $\beta$ )									
17		0.04	=BETADIST(x, $\alpha$ , $\beta$ )									

Sheet1 Sheet1 (2) Sheet2 Sheet3

Ready 147%

# THE PRIORS

$$\pi = \text{Beta}[D_P + D_N + \alpha_{\text{prior}}, (T_P - D_P) + (T_N - D_N) + \beta_{\text{prior}}]$$

$$Se = \text{Beta}(D_P + \alpha_{\text{prior}}, D_N + \beta_{\text{prior}})$$

$$Sp = \text{Beta}[(T_N - D_N) + \alpha_{\text{prior}}, (T_P - D_P) + \beta_{\text{prior}}]$$

- Now we have the data to put in our equations



Now we  
have the  
basic  
know-  
ledge and  
we can  
make  
things more complex

