

# Distributional Probability Logic

**Gernot D. Kleiter**

Department of Psychology  
University of Salzburg, Austria

Supported by the Austrian Science Foundation (I00141)  
within the LogICCC programme of the European Science Foundation  
LcpR.

1st idea

Distributional  
logic

Second order  
probabilities

Probability  
distributions

Probability  
logic

First order  
Probabilities

Point/interval  
probabilities

Propositional  
logic

Propositions  
events

Truth values

2nd idea

Probability  
logic

Conditional  
IF-THEN

Combine

Conditional  
events

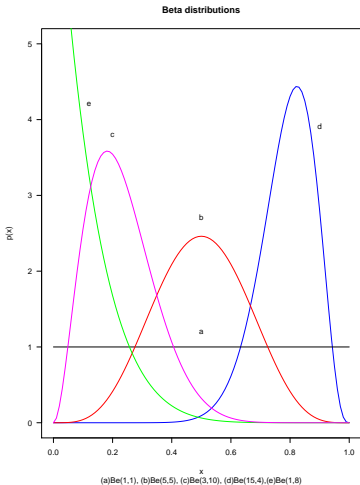
Probability  
distributions

Imprecise  
probabilities

Undetermined  
truth value

# Beta distributions

- ▶ X-axis: First order probability
- ▶ Y-axis: Second order probability density
- ▶ Area under the curve: Second order probability



# Beta distributions

are a family of probability density functions on the unit interval  $[0, 1]$ .

## Definition (Standard beta distribution)

A random variable  $X$  is standard beta distributed with shape parameters  $\alpha$  and  $\beta$  if

$$p(x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}. \quad (1)$$

We write for short  $X \sim \text{Be}(\alpha, \beta)$  or  $\text{Be}(x; \alpha, \beta)$ .

Mean and variance

$$E(X) = \frac{\alpha}{\alpha + \beta} \quad (2)$$

$$\text{Var}(X) = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}. \quad (3)$$

$X$  is a first order probability;  $p(x)$  is a second order pdf. The shape parameters  $\alpha$  and  $\beta$  are related to the frequencies of favorable and non-favorable cases.

$\text{Be}(1, 1)$  is the uniform distribution. If  $\alpha > 1$  and  $\beta > 1$  the distributions is unimodal, if either  $\alpha < 1$  or  $\beta < 1$  it is J-shaped, and if  $\alpha < 1$  and  $\beta < 1$  it is U-shaped.

# Elementary operators

**Negation** . The shape parameters of the beta distribution switch positions: We write  $P(A) = X$  and  $P(\neg A) = 1 - X$ .

If  $X \sim Be(a, b)$ , then  $1 - X \sim Be(b, a)$ .

**Conjunction** Linda task ...

## Linda task (Tversky & Kahneman, 1983)

Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.

Rank order the probabilities that

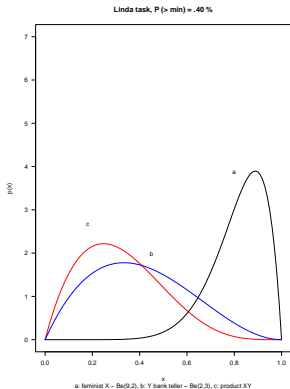
- ▶ Linda is a bank teller.  $X \sim Be(2, 3)$
- ▶ Linda is active in the feminist movement.  $Y \sim Be(9, 2)$
- ▶ Linda is a bank teller and is active in the feminist movement.  
 $W = XY \sim \dots?$

$X$  and  $Y$  are not assumed to be independent!

What is the probability that  $w > x$ ? For one value of  $w$  the probability that  $w$  exceeds  $p(x)$  is  $p(w)F(w; a, b)$ , where

$$P(w > x) = \int_0^1 f(w)F(w > x; a, b)dw \quad (4)$$

In the example:  $P(w > x) = .40$ .



Obtaining the conjunction from the product

$$p(x, z) = Be(x; a, b)Be(z; e, f). \quad (5)$$

We change the metric  $x \mapsto w$ , where  $w = xz$  and  $x = w/z$ . Jacobian (positive first derivative) is  $|dx/dw| = 1/z$  so that

$$p(w, z) = \frac{1}{z} Be\left(\frac{w}{z}; a, b\right) Be(z; e, f). \quad (6)$$

The marginal  $p(w)$  is obtained by integrating out  $z$ ,

$$p(w) = \int_0^1 \frac{1}{z} Be\left(\frac{w}{z}; a, b\right) Be(z; e, f) dz. \quad (7)$$

General Case: Obtaining the conjunction from bivariate  
If  $x$  and  $y$  are fixed, then  $w$  is between a lower and an upper bound,

$$\max\{0, x + y - 1\} \leq w \leq \min\{x, y\}$$

If we know “nothing” about the correlation  $\Rightarrow$  rectangular distribution.  
Considering all points  $(x, y)$  in the plane  $[0, 1]^2$  we get two hill-like surfaces above, one for the lower bound and one for the upper bound.  
We marginalize with respect to  $w$  by collecting for every  $w$  the the weighted average of the number of times  $w$  is in the interval between the two surfaces. The weights are equal to  $p(x, y)$ .  
 $p(x, y)$  is bivariate!

# Bivariate beta distributions

There are many proposals, first ones going back to Karl Pearson, 1923, the latest and best proposal by Nadarajah, 2007

Figure: Truncated bivariate beta distribution for  $p_1 = 2, p_2 = 4, p_3 = 10$

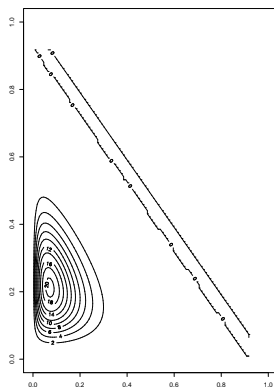


Figure: Truncated bivariate beta distribution for  $p_1 = 10, p_2 = 4, p_3 = 10$

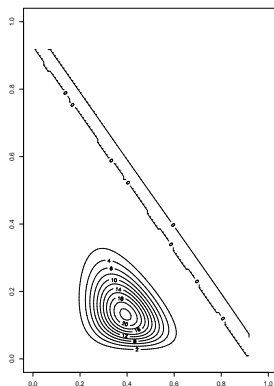


Figure: Bivariate beta distribution for  $\alpha = 2, \beta = 4, \gamma = 8, \delta = .8$

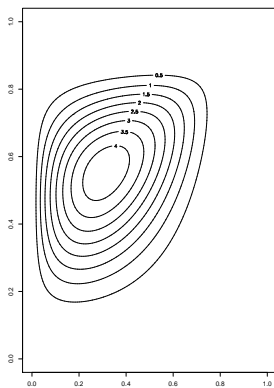


Figure: Bivariate beta distribution for  $\alpha = 2, \beta = 2, \gamma = 8, \delta = 0$

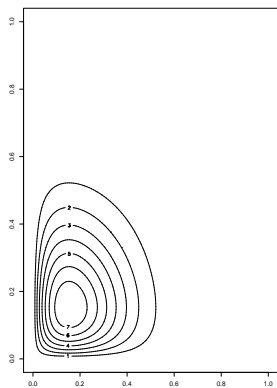


Figure: Bivariate beta distribution for  $\alpha = 3, \beta = 4, \gamma = 6, \delta = .3$

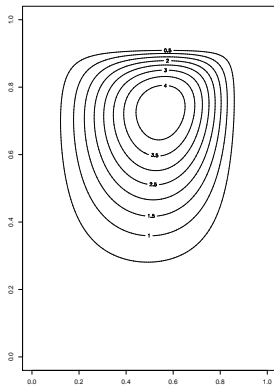
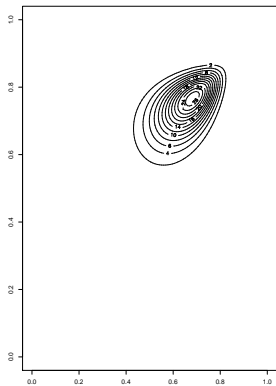


Figure: Bivariate beta distribution for  $\alpha = 10, \beta = 15, \gamma = 25, \delta = .9$



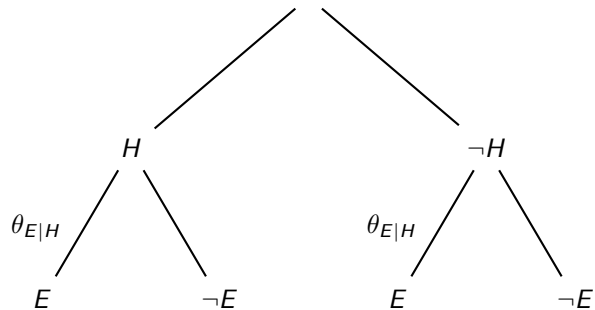
# The conditional as a conditional event

$$B|A = \begin{cases} 1 & \text{if } A = 1 \quad B = 1 \quad \text{win} \\ 0 & \text{if } A = 1 \quad B = 0 \quad \text{lose} \\ ? & \text{if } A = 0 \quad \text{money back} \end{cases} \quad (8)$$

Representation in the distributional approach

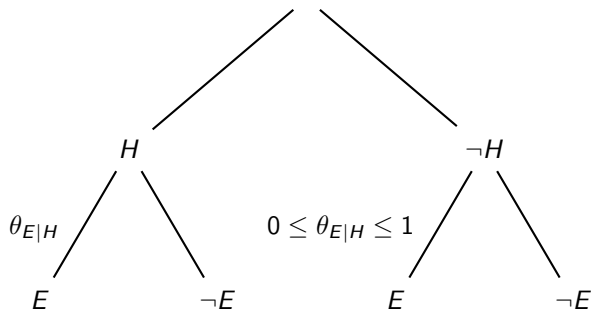
$$p(B|A) \sim \begin{cases} \text{Be}(a, b) & \text{if } A = 1 \quad B = 1 \\ \text{Be}(b, a) & \text{if } A = 1 \quad B = 0 \\ \text{Be}(1, 1) & \text{if } A = 0 \end{cases} \quad (9)$$

# Probabilistic representation of the conditional

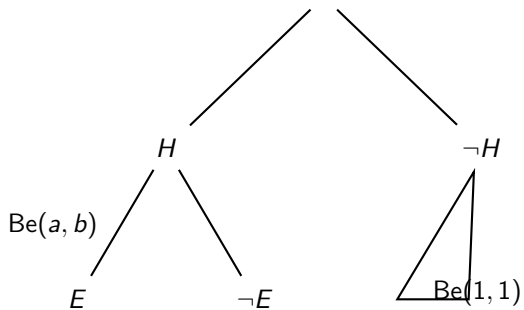


# Probabilistic representation of the conditional

The “?” becomes a “[0,1] interval”



The  $[0,1]$  interval becomes a uniform distribution



# Forward inference, MP & DA

In standard logic the MODUS PONENS has the argument form

$$\frac{A \rightarrow B \quad A}{B}$$

In the distributional approach

$$P(B|A) = U \sim Be(a, b)$$

$$P(B|\neg A) = V \sim Be(c, d)$$

$$P(A) = X \sim Be(g, h)$$

---

$$P(B) = Y = XU + (1 - X)V$$

with  $V \sim Be(1, 1)$  if antecedent is “void”

$$p(x, u, v) = Be(x; g, h)Be(u; a, b)Be(v; c, d) \quad (10)$$

change the metric of  $V$  to  $Y$ ,  $V \mapsto Y$  with the linear transformation

$$y = xu + (1 - x)v. \quad (11)$$

The Jacobian is

$$|dy/dv| = 1 - x. \quad (12)$$

The inverse of the transformation is

$$v = \frac{y - xu}{1 - x}. \quad (13)$$

Substitute the value of the inverse of  $y$  and multiply with the Jacobian

$$p(x, u, y) = Be(x; g, h)Be(u; a, b)Be\left(\frac{y - xu}{1 - x}; c, d\right)(1 - x). \quad (14)$$

integrate out  $x$  and  $u$

$$p(y) = \int_{x=0}^{x=1} \int_{u=u_1}^{u=u_2} Be(x; g, h)Be(u; a, b)Be\left(\frac{y - xu}{1 - x}; c, d\right)(1 - x)dxdu, \quad (15)$$

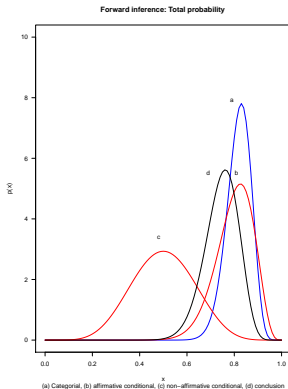
Find the lower and upper bounds of the inner integral we take the extreme values for  $v$  in

$$u = \frac{y - (1 - x)v}{x}. \quad (16)$$

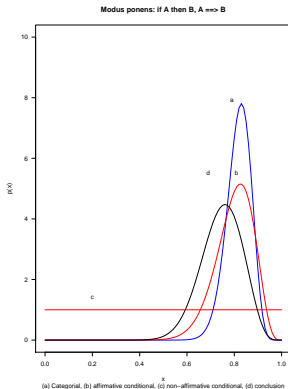
If  $v = 0$ , then we have  $u = y/x$  and for  $v = 1$  we have  $u = \frac{y+x-1}{x}$ .

$$u_1 = \max\left\{0, \frac{y+x-1}{x}\right\} \quad \text{and} \quad u_2 = \min\left\{1, \frac{y}{x}\right\} \quad (17)$$

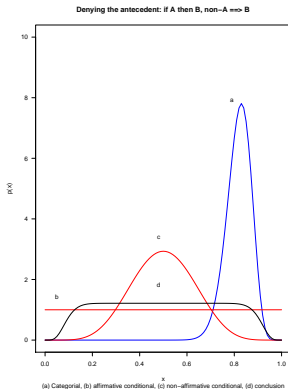
**Figure:** Total probability. Base rate  $Be(45,10)$ , affirmative  $Be(20,5)$ , non-affirmative  $Be(7,7)$ .  $h(\text{Base rate}) = -2.26$ ,  $h(\text{affirmative}) = -1.67$ ,  $h(\text{non-affirmative}) = -0.91$ ,  $h(\text{conclusion}) = -1.79$ .



**Figure:** Modus ponens. Base rate  $Be(45,10)$ , affirmative  $Be(20,5)$ , non-affirmative  $Be(1,1)$ .  $h(\text{Base rate}) = -2.26$ ,  $h(\text{affirmative}) = -1.67$ ,  $h(\text{non-affirmative})=0$ ,  $h(\text{conclusion})=-1.50$ .



**Figure:** A form of denying the antecedent. Base rate  $Be(45,10)$ , affirmative  $Be(20,5)$ , non-affirmative  $Be(1,1)$ .  $h(\text{Base rate}) = -2.26$ ,  $h(\text{affirmative}) = 0$ ,  $h(\text{non-affirmative}) = -0.91$ ,  $h(\text{conclusion}) = -0.18$ .



# Backward inference, MT & AC

The argument form of the MT is

$$\frac{A \rightarrow B \quad \neg B}{\neg A}$$

$$P(B|A) = U \sim \text{Be}(a, b)$$

$$P(B|\neg A) = V \sim \text{Be}(c, d)$$

$$P(\neg B) = 1 - Y \sim \text{Be}(h, g)$$

$$P(A) = X = \frac{Y - V}{U - V}$$

**Table:** Backward inference. Note that  $Y \sim \text{Be}(g, h)$  and its the “negation”  $1 - Y$  is  $\text{Be}(h, g)$ .

We write down the joint distribution of  $(Y, U, V)$

$$p(y, u, v) = Be(y; h, g)Be(u; a, b)Be(v; c, d) \quad (18)$$

We focus on the third variable  $V$  and change its metric  $V \mapsto X$  with the transformation

$$x = \frac{y - v}{u - v}. \quad (19)$$

The Jacobian is

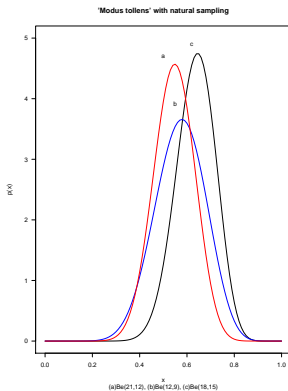
$$\left| \frac{dx}{dv} \right| = \left| \frac{-1}{u - v} \right| \quad (20)$$

and the inverse of the transformation is

$$v = \frac{y - xu}{1 - x}. \quad (21)$$

... *Idotsldots* Analog to MP, but watch the denominator in the Jacobian!

Figure: “Modus tollens” with natural sampling



# Bayes' Theorem

$$P(A|B) = \frac{P(A)P(B|A)}{P(A)P(B|A) + (1 - P(A))P(B|\neg A)}. \quad (22)$$

$$z = \frac{xu}{xu + (1 - x)v}. \quad (23)$$

We assume  $X$ ,  $U$  and  $V$  to be independent and express the joint distribution as the product of three beta distributions

$$p(x, u, v) = Be(x; g, h)Be(u; a, b)Be(v; c, d). \quad (24)$$

We change the metric of  $V \mapsto Z$

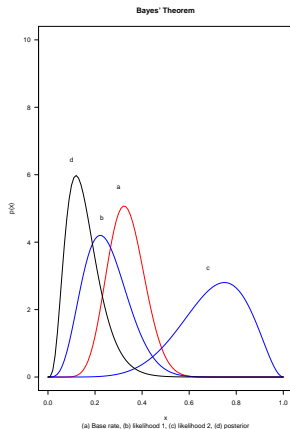
$$v = \frac{xu(1 - z)}{(1 - x)z}. \quad (25)$$

The Jacobian is obtained by the quotient rule and is

$$\left| \frac{dv}{dz} \right| = \left| \frac{xu}{(1 - x)z^2} \right| \quad (26)$$

..... Analog steps as above

Figure: Bayes Theorem. Base rate  $Be(12, 24)$ , likelihood 1  $Be(5, 15)$ , likelihood 2  $Be(7, 3)$ .



1. Chaining, propagation of imprecision
2. Approximations: there are many good ones
3. Mixtures, Dirichlet distributions
4. U-shaped betas
5. Entropy, argument strength
6. Copulas, connection to t-norms