# Bayesian analysis of hypotheses

- **Hypotheses:**
  - H could be about a parameter: ”$\theta$<0”
  - Compute $P(H_0|X)=P(\theta < 0 \,|X)$, the cumulative posterior density at 0. Alternative hypothesis $H_1$ = ”$\theta$>=0”
  - $P(H_0 \mid X)$ and $P(H_1 \mid X)$ possible to compute if ”H” is a region of parameter space.
  - Can use indicator variables in BUGS to compute these probabilities.
  - In contrast to classical hypothesis testing: this does not reject or accept a H. Just calculate the probability of H, given evidence.

# Two sample problems

- **Comparison of two populations:**
  - Prevalence comparisons in population studies. (Binomial proportions)
    - Also e.g. gallup studies (Multinomial proportions)
  - Compare treatment group vs control group in medical problems, case-control studies in epidemiology, for example:
    - Effect of new medicine? - treatment vs control.
    - Does vaccine cause autism? - case-control study.
  - Compare measurements between groups (normal models)

# Two sample problems

- **Two binomial proportions**
  - $X_1 \sim \text{Bin}(p_1,n_1)$
  - $X_2 \sim \text{Bin}(p_2,n_2)$
    - Question: $p_1 > p_2$ ? **= This is hypothesis**
  - Compute: $\pi(p_1 , p_2 \mid X_1, X_2 , n_1 , n_2 )$
    - Populations assumed to be separate.
    - Samples assumed to be independent conditionally, given $p_1$, $p_2$.
    - Independent priors assumed: $\pi(p_1 , p_2) = \pi(p_1)\, \pi(p_2)$.

# Two sample problems

- **Two binomial proportions**.
  - It's enough to compute two posterior distributions:

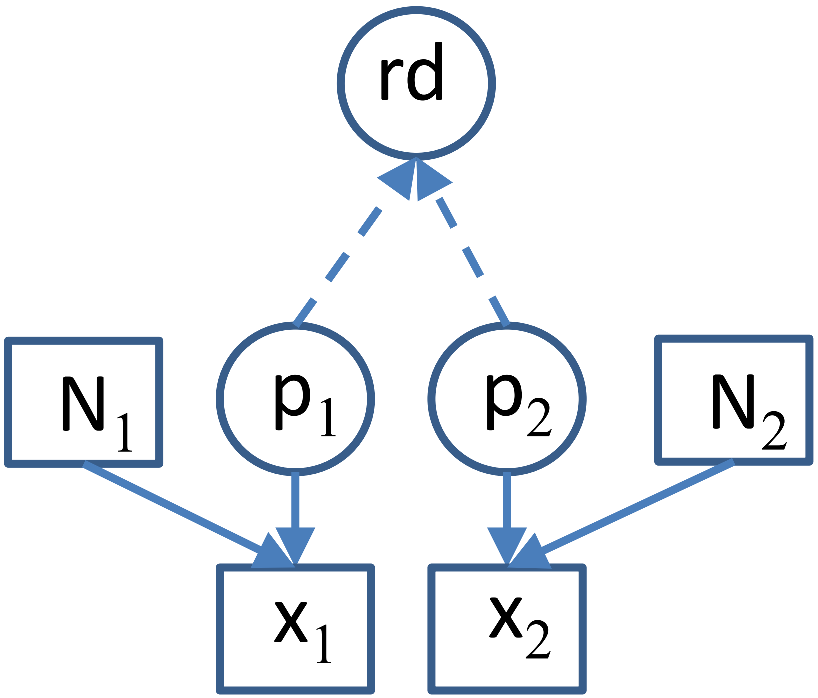
$\pi(p_1,p_2 \mid X_1,X_2,n_1,n_2) = \pi(X_1,X_2 | n_1,n_2,p_1,p_2)\pi(p_1)\pi(p_2)/c$
$= \pi(X_1 | n_1,p_1)\pi(p_1)\ \pi(X_2 | n_2,p_2)\pi(p_2)/c$

  - Draw Monte Carlo sample from the two posteriors (if conjugate priors, these are beta-densities)
  - Compute approximately $P(p_1 > p_2)$ from the sample.
  - Could also compute posterior of :
    - 'risk ratio' $rr = p_1/p_2$
    - 'risk difference' $rd = p_1 - p_2$
    - 'odds ratio' $or = [\ p_1/(1-p_1)\ ]\ /\ [\ p_2/(1-p_2)\ ]$

# Two sample problems

- **In BUGS:**

```
model{
for(i in 1:2){
x[i]  ~ dbin(p[i],n[i])
p[i] ~  dbeta(1,1)
}
P <-  step(p[1]-p[2])
rr <- p[1]/p[2];   rd <- p[1]-p[2]
}
list(x=c(,),n=c(,))
```



# Bayes factors

- Sometimes used: **posterior odds =** $P(H_0 \mid X)/P(H_1 \mid X)$.
  - If ”>1”, shows support for $H_0$.
  - **Bayes factor: a ratio of posterior and prior odds**

  - $BF = [\ P(H_0 \mid X)/P(H_1 \mid X)\ ] \ / \ [\ P(H_0)/P(H_1)\ ]$

    $= [P(H_0 \mid X)\ P(H_1)\ ] \ / \ [P(H_1 \mid X)\ P(H_0)]$

**Posterior odds = Prior odds x BF**

This is a different way of expressing Bayes theorem:
BF expresses how much data change the prior odds.

# Simple point hypothesis

- A point hypothesis $H_0 : \theta = \theta_0$ against $H_1 : \theta = \theta_1$
  - We must have positive probability $P(H_0)=1-P(H_1)$
  - BF then becomes the same as **’likelihood ratio’ (uskottavuusosamäärä)**

$$\frac{\pi(\theta=\theta_0 \mid X)}{\pi(\theta=\theta_1 \mid X)} = \frac{\pi(\theta=\theta_0)}{\pi(\theta=\theta_1)} \boxed{\frac{\pi(X \mid \theta=\theta_0)}{\pi(X \mid \theta=\theta_1)}}$$

Because constant $\pi(X)$ cancels out.

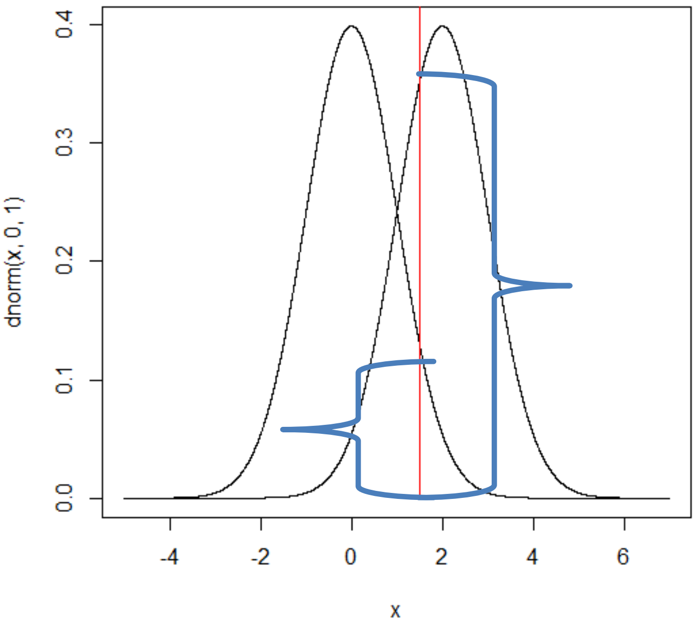
- But: how big (small) BF is big (small) enough ?
- Composite hypothesis, one-sided, two-sided…

# Point hypothesis example

- **Assume X ~ N(θ,1), data: X=1.5**
  - A point hypothesis $H_0 : \theta = 0$ against $H_1 : \theta = 2$.
  - Assume prior $\pi(\theta=0)=\pi(\theta=2)=0.5$
  - Then, posterior odds = likelihood ratio.
  - Conversion to probability :

    p = odds/(1+odds)

$$\frac{\pi(\theta=\theta_0 \mid X)}{\pi(\theta=\theta_1 \mid X)} = \frac{\pi(\theta=\theta_0)}{\pi(\theta=\theta_1)} \boxed{\frac{\pi(X \mid \theta=\theta_0)}{\pi(X \mid \theta=\theta_1)}}$$



# Composite hypothesis

- Composite hypothesis, one-sided: $H_0 : \theta < \theta_0$ against $H_1 : \theta \geq \theta_0$
  - Bayes-factor is

$$BF = \frac{\int_{\Theta_0} \pi(\theta \mid X) d\theta \int_{\Theta_1} \pi(\theta) d\theta}{\int_{\Theta_1} \pi(\theta \mid X) d\theta \int_{\Theta_0} \pi(\theta) d\theta} = \frac{\int_{\Theta_0} \pi(X \mid \theta) \pi(\theta) d\theta \int_{\Theta_1} \pi(\theta) d\theta}{\int_{\Theta_1} \pi(X \mid \theta) \pi(\theta) d\theta \int_{\Theta_0} \pi(\theta) d\theta}$$

  which now depends also on prior.

  - Two-sided hypothesis $H_0 : \theta = \theta_0$ against $H_1 : \theta \neq \theta_0$ does not make sense unless we set positive prior probability for $H_0$.

# Composite hypothesis

- Example: $H_0 : \theta = \theta_0$ against $H_1 : \theta \neq \theta_0$
  - Assume a coin is either fair so that P(heads)=P(tails)=0.5= $\theta_0$, or the coin is biased so that P(heads) = $\theta$, P(tails)=1-$\theta$.
  - Parameter $\theta$ is unknown. Without much prior information, we choose the prior $\pi(\theta)$=U(0,1).
  - **The model for observed data is x~Bin($\theta$,n)**
  - The hypothesis is about two possible models:
  - $M_0$ with $\theta = \theta_0 = 0.5$ describing fair coin
  - $M_1$ with $\theta \sim U(0,1)$ describing unknown bias

# $M_0$ or $M_1$?

- Bayes factor is comparing the two models
- BF = posterior odds / prior odds

$$BF = \frac{P(M_0 \mid X) / P(M_1 \mid X)}{P(M_0) / P(M_1)}$$

$$= \frac{P(M_0)P(X \mid M_0) / (P(M_1)P(X \mid M_1))}{P(M_0) / P(M_1)} = \frac{P(X \mid M_0)}{P(X \mid M_1)}$$

$$BF = \frac{P(X \mid \theta = 0.5)}{\int_0^1 P(X \mid \theta)\pi(\theta \mid M_1)d\theta}$$

# $M_0$ or $M_1$?

- Assume X=115, N=200

$$P(X=115 \mid M_0) = \binom{200}{115} 0.5^{200} = 0.005956$$

$$P(X=115 \mid M_1) = \int_0^1 \binom{200}{115} \theta^{115} (1-\theta)^{85} d\theta = \frac{1}{201} = 0.004975$$

- BF=1.197, slightly supporting $M_0$.
- If priors $P(M_0)=P(M_1)$, then posterior odds = BF, and $P(M_0|X)$= odds/(1+odds)≈1/(1+1.197)≈0.54