Dec 14

Laplace's Prior: the prior of indifference / uniformity.
In the Possion model θ in (0, infinity), we need a distribution that is uniform on that set. A distribution would look like:

$$P(\theta) = c > 0, \int_0^{\infty} P(\theta)\,d\theta = \int_0^{\infty} c\,d\theta = c\int_0^{\infty} 1\,d\theta = \infty \Rightarrow P(\theta) = c \quad \text{d.n.e.}$$

there cannot be a proper Laplace prior. But there is an improper Laplace prior.

$P(\theta) \propto 1 \Rightarrow$ Laplace's idea

$$P(\theta|x) \propto P(x|\theta)P(\theta) = e^{-n}\theta^{\Sigma x_i} P(\theta) \propto e^{-n\theta}\theta^{\Sigma x_i + 1 - 1}$$

$$\propto \text{Gamma}(\underbrace{1 + \Sigma x_i}_{\alpha}, \underbrace{n}_{0+n-\beta})$$

$\Rightarrow P(\theta) = \text{Gamma}(1, 0)$, an improper prior

$\Downarrow$ implies

$X_0 = 1, n_0 = 0$ nonsense!

Is the posterior proper? Yes. ALWAYS. Since sum $X_i >= 0$, it's first parameter is always $>= 1 > 0$ and since $n >= 1$, it's second $\rightarrow$

parameter is always >= Since sum $X_i >= 0$, it's
first parameter is always $>= 1 > 0$ and since $n >=$
1 it's second parameter.

Haladane prior of comple ignorance, Setting all
pseudodata to zero i.e. $X_0 = 0, n_0 = 0 \Rightarrow P(\theta) = \text{Gamma}(0,0)$

improper

$\Rightarrow P(\theta|x) = \text{Gamma}(\sum x_i, n) = \hat{\theta}_{MMSE} = \frac{\sum x_i}{n} = \bar{x} = \hat{\theta}_{MLE}$

Is this posterior proper? only if sum $x_i > 0$.

Jeffrey's prior. $P_J(\theta) \propto \sqrt{I(\theta)} = \sqrt{\frac{n}{\theta}} \propto \theta^{-\frac{1}{2}} \propto \text{Gamma}(\frac{1}{2}, 0)$ improper

$P(\theta|x) = \text{Gamma}(\frac{1}{2} + \sum x_i, 0 + n)$

Is it proper? Yes Always!

$\ldots l'(\theta) = -n + \frac{\sum x_i}{\theta} \Rightarrow l''(\theta) = -\frac{\sum x_i}{\theta^2}$

$I[\theta] = E_X[-l''(\theta)] = \frac{E[\sum x_i]}{\theta^2} = \frac{n E[X_i]}{\theta^2} = \frac{n}{\theta}$

$X \sim \text{Poission}(\theta)$

$E[X] = \sum_{x \in supp(X)} x \, P(x) = \sum_{x=0}^{\infty} x \frac{e^{-\theta}\theta^x}{x!} = e^{-\theta} \sum_{x=1}^{\infty} \frac{x\theta^x}{x!} = e^{-\theta} \sum_{x=1}^{\infty} \frac{\theta^x}{(x-1)!}$

$y = x - 1 \Leftrightarrow x = y + 1$

$e^{-\theta} \sum_{y=0}^{\infty} \frac{\theta^{y+1}}{y!} = \theta e^{-\theta} \underbrace{\sum_{y=0}^{\infty} \frac{\theta^y}{y!}}_{e^{\theta}} = \theta \cancel{e^{-\theta} e^{\theta}} = \theta$

stribution
poission
of where

Posterior predictive distribution. You see n observations and you want the distribution of n^* future observations is

For our case here, we let $n^* = 1$.

$x_1, x_2, \ldots, x_n \sim$ iid poission(θ) | $X_* \sim$?

n | $n_* = 1$

← past → | ← future →

$$P(X_*|X) = \int_{\Theta} \underbrace{P(X_*|\theta)}_{\sim \text{Poission}(\theta)} \underbrace{P(\theta|X)}_{\sim \text{Gamma}(\alpha+\Sigma x_i, \beta+n)} d\theta$$

what is the support of Θ

$$= \int_0^\infty \frac{e^{-\theta}\theta^{x_*}}{x_*!} \frac{(\beta+n)^{\alpha+\Sigma x_i}}{\Gamma(\alpha+\Sigma x_i)} \theta^{\alpha+\Sigma x_i-1} e^{-(\beta+n)\theta} d\theta$$

$$= \frac{(\beta+n)^{\alpha+\Sigma x_i}}{x_*!\,\Gamma(\alpha+\Sigma x_i)} \int_0^\infty e^{-\theta}\theta^{x_*}\theta^{\alpha+\Sigma x_i-1} e^{-(\beta+n)\theta} d\theta$$

$$= \frac{(\beta+n)^{\alpha+\Sigma x_i}}{x_*!\,\Gamma(\alpha+\Sigma x_i)} \int_0^\infty \theta^{x_*+\alpha+\Sigma x_i-1} e^{-(\beta+n+1)\theta} d\theta$$

let $t = (\beta+n+1)\theta \Rightarrow \theta = \frac{t}{\beta+n+1} \Rightarrow \frac{d\theta}{dt} = \frac{1}{\beta+n+1} \Rightarrow d\theta = \frac{1}{\beta+n+1} dt$

if $\theta = 0 \Rightarrow t = 0$; if $\theta = \infty \Rightarrow t = \infty$

$$\frac{(\beta+n)^{\alpha+\Sigma x_i}}{x_*!\,\Gamma(\alpha+\Sigma x_i)} \int_0^\infty \frac{t^{x_*+\alpha+\Sigma x_i-1}}{(\beta+n+1)^{\cancel{x_*+\alpha+\Sigma x_i-1}}} e^{-t} \frac{1}{\cancel{\beta+n+1}} dt \rightarrow$$

$$= \frac{(\beta+n)^{\alpha+\Sigma x_i}}{x_*!\,\Gamma(\alpha+\Sigma x_i)(\beta+n+1)^{x_\alpha+\alpha+\Sigma x_i}} \underbrace{\int_0^\infty t^{(x^*+\alpha+\Sigma x_i)-1} e^{-t}\,dt}_{\text{gamma integral}}$$

$$= \frac{(\beta+n)^{\alpha+\Sigma x_i}\,\Gamma(x_\alpha+\alpha+\Sigma x_i)}{x_*!\,\Gamma(\alpha+\Sigma_{x_i})(\beta+n+1)^{x_\alpha+\alpha+\Sigma x_i}}$$

$$= \frac{(\beta+n)^{\alpha+\Sigma x_i}}{(\beta+n+1)^{\alpha+\Sigma x_i}} \frac{1}{(\beta+n+1)^{x_*}} \frac{\Gamma(x_*+\alpha+\Sigma x_i)}{x_*!\,\Gamma(\alpha+\Sigma x_i)}$$

$$= \left(\frac{\beta+n}{\beta+n+1}\right)^{\alpha+\Sigma x_i} \left(\frac{1}{\beta+n+1}\right)^{x_*} \frac{\Gamma(x_*+\alpha+\Sigma x_i)}{x_*!\,\Gamma(\alpha+\Sigma x_i)}$$

let: $= \frac{\beta+n}{\beta+n+1} \in (0,1)$, $1-p = \frac{1}{\beta+n+1} \in (0,1)$, $r := \Sigma x_i + \alpha > 0$

$$\Downarrow$$

$$= p^r (1-p)^{x_*} \frac{\Gamma(x_*+r)}{x_*!\,\Gamma(r)} = \text{Ext NegBin}(r,p)$$

extended negative binomial random variable model

if $\alpha \in \{0,1,2,\ldots\}$

$$\Downarrow$$

$$= \binom{x_*+r-1}{r} p^r (1-p)^{\alpha_*} = \text{NegBin}(r,p)$$

Thus the posterior predictive distribution has variance up to 2x the poisson (i.e more spread out or less sure of where the realization will be.)

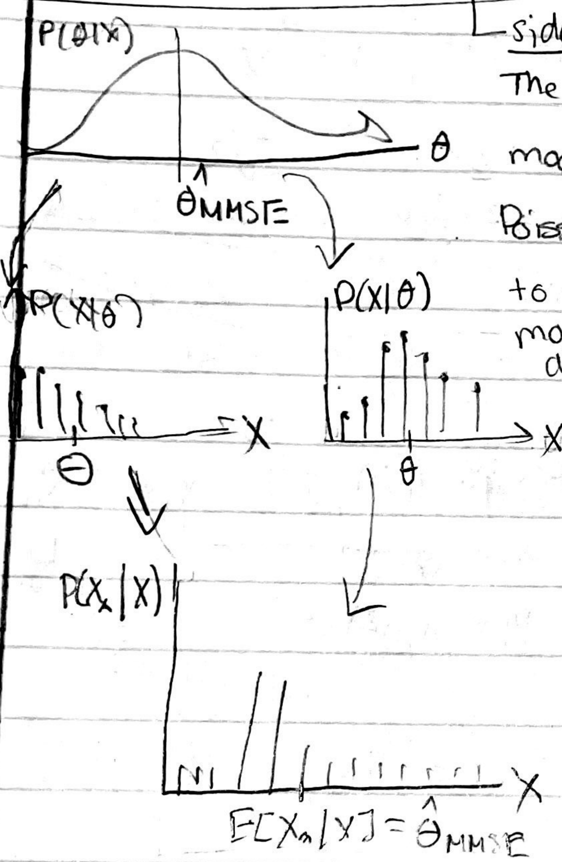
From 368... the negative binomial is the sum of iid geometric random variable. Since the expectation of the geometric rv is $(1-p)/p$, the expectation of the negative binomial by linearity is

$$P(X_*|X) = \text{ExtNegBin}(r,p) \Rightarrow E[X_*|X] = r\,\frac{1-p}{p}$$

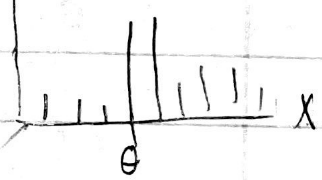
$$= (\Sigma x_i + \alpha)\,\frac{\frac{1}{n+\beta+1}}{\frac{n+\beta}{n+\beta+1}}$$

$$= \frac{\Sigma x_i + \alpha}{n+\beta} = \hat{\theta}_{MMSE}$$

lec15.



side Note:
The negative binomial model is a "dispersed Poisson" analagous to the betabinomial model being a dispersed binomial



$$Var[X_*|X] = r\,\frac{1-p}{p^2} = \hat{\theta}_{MMSE}\cdot\frac{1}{p} = \frac{\beta+n+1}{\beta+n}$$

$$= \hat{\theta}_{MMSE}$$

Curly F: one $(N\,\theta, \sigma^2) = P(x|\theta,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\theta)^2}$

let σ^2 be fixed/known in advance

$$P(X|\theta,\sigma^2) \propto e^{-\frac{1}{2\sigma^2}(x-\theta)^2}$$

$$= e^{-\frac{1}{2\sigma^2}(x^2 - 2x\theta + \theta^2)}$$