Lec 19 Math 341 3/28/18

$X_1, \dots X_n \overset{iid}{\sim} \text{Poisson}(\theta)$ From last class.. $\hat{\theta}_{MLE} = \bar{X} = \frac{\sum x_i}{n}$

$\theta \sim \text{Gamma}(\alpha, \beta)$

$\Rightarrow \theta | X_1, \dots, X_n \sim \text{Gamma}(\underbrace{\sum x_i + \alpha}_{\alpha'}, \underbrace{n + \beta}_{\beta'})$

Pseudocount interpretation?

α = # prior successes

β = # prior trials

$E(\theta) = \frac{\alpha}{\beta}$ = prior avg # success per # trials

Prior expectation

Uninformative Priors

$P(\theta) \propto 1 \Rightarrow P(\theta) = \text{Gamma}(1, 0) \Rightarrow \theta | x \sim \text{Gamma}(\sum x_i + 1, n)$ Indifference

Seeing 1 success in zero trials (weird)

never improper

$P(\theta) = \text{Gamma}(0, 0) \Rightarrow \theta | x \sim \text{Gamma}(\sum x_i, n)$

(Haldane)

improper if no successes

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

never improper

Shrinkage? Yes, standard weighted arithmetic avg shrinkage

$$\hat{\theta}_{MMSE} = \frac{\sum x_i + \alpha}{n+\beta} = \frac{n}{n}\cdot\frac{\sum x_i}{n+\beta} + \frac{\alpha}{n+\beta}\cdot\frac{\beta}{\beta} = \overbrace{\frac{n}{n+\beta}}^{1-\rho}\,\overbrace{\frac{\sum x_i}{n}}^{\hat{\theta}_{MLE}} + \overbrace{\frac{\beta}{n+\beta}}^{\rho}\,\overbrace{\frac{\alpha}{\beta}}^{E(\theta)}$$

$$\lim_{n\to\infty} \rho = 0 \checkmark$$

Posterior pred. distr.

$X|\theta \sim \text{Poisson}(\theta)$

Now θ is unknown ... so this will be "dispersed"

the ~~beta-binomial~~ is the "dispersed binomial"

the beta geometric is the "dispersed geometric"

what is a "dispersed Poisson"?

$$P(X^*|X) = \int_{\Theta} P(X^*|\theta)\, P(\theta|X)\, d\theta \qquad \text{for } n^* = 1 \ldots$$

$$= \int_0^\infty \left(\frac{e^{-\theta}\,\theta^{x^*}}{x^*!}\right)\left(\frac{\beta'^{\alpha'}}{\Gamma(\alpha')}\, e^{-\beta'\theta}\,\theta^{\alpha'-1}\right) d\theta$$

$$= \frac{\beta'^{\alpha'}}{\Gamma(\alpha')\, x^*!}\int_0^\infty \underbrace{e^{-(\beta'+1)\theta}\,\theta^{x^*+\alpha'-1}}_{\text{kernel of } \text{Gamma}(x^*+\alpha',\, \beta'+1) = \frac{1}{c}(k(\theta))} d\theta = \frac{\beta'^{\alpha'}}{\Gamma(\alpha')\, x^*!}\cdot c \cdot \int_0^\infty \underbrace{\frac{1}{c}\, k(\theta)}_{P(\theta)}\, d\theta$$

$$= \frac{\beta'^{\alpha'}}{\Gamma(\alpha')\, x^*!} \frac{\Gamma(x^*+\alpha')}{(\beta'+1)^{x^*+\alpha'}} \int_0^\infty \underbrace{\frac{(\beta'+1)^{x^*+\alpha'}}{\Gamma(x^*+\alpha')} e^{-(\beta'+1)\theta}\, \theta^{x^*+\alpha'-1}\, d\theta}_{\substack{\Downarrow \text{ PDF of Gamma}(x^*+\alpha',\ \beta'+1) \\ 1}}$$

$$= \left(\frac{\beta'}{(\beta'+1)}\right)^{\alpha'} \left(\frac{1}{\beta'+1}\right)^{x^*} \frac{\Gamma(x^*+\alpha')}{x^*!\,\Gamma(\alpha')}$$

Note: $\frac{\beta'}{\beta'+1} \in (0,1)$ let $p := \frac{\beta'}{\beta'+1} \Rightarrow 1-p = \frac{1}{\beta'+1}$

$$= \frac{\Gamma(x^*+\alpha')}{x^*\,\Gamma(\alpha')} (1-p)^{x^*} p^{\alpha'} = \text{ExtNegBin}(\alpha', p)$$ "fractional successes"

If $\alpha' \in \mathbb{N}$ note $x^* \in \mathbb{N}$ so... $\Gamma(n) = (n-1)!$

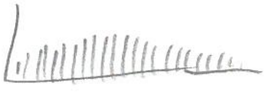
$$= \frac{(x^*+\alpha'-1)!}{x^*!\,(\alpha'-1)!} (1-p)^{x^*} p^{\alpha'} = \binom{x^*+\alpha'-1}{x^*} (1-p)^{x^*} p^{\alpha'} = \text{NegBin}(\alpha', p)$$

# of failures before α' successes
where each experim $\overset{iid}{\sim}$ Bern(p)

Poisson dispersal is a neg-binomial! $\cancel{E[x^*|x] = \frac{pr}{1-p} = \frac{\frac{n+\beta}{n+\beta+1}(\sum x_i+\alpha)}{\frac{1}{n+\beta+1}} = (n+\beta)(\sum x_i+\alpha)}$

$\cancel{Var[x^*|x] = \frac{pr}{(1-p)^2}}$ $\cancel{\frac{\left(\frac{n+\beta}{n+\beta+1}\right)(\sum x_i+\alpha)}{\frac{1}{(n+\beta+1)^2}}}$

$$\Rightarrow x^*|X \sim \text{NegBin}\left(\alpha', \frac{\beta'}{\beta'+1}\right) = \text{NegBin}\left(\sum x_i + \alpha,\ \frac{n+\beta}{n+\beta+1}\right)$$



[illegible]

Normal model

Normal Model

$$P(X \mid \theta, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(X-\theta)^2}$$

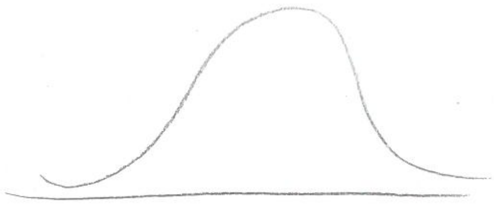
$$\propto e^{-\frac{1}{2\sigma^2}(x^2 - 2\theta x + \theta^2)}$$

$$= e^{-\frac{x^2}{2\sigma^2}} e^{\frac{\theta x}{2\sigma^2}} e^{-\frac{\theta^2}{2\sigma^2}}$$

$$\propto e^{-\frac{x^2}{2\sigma^2}} e^{\frac{\theta x}{2\sigma^2}}$$

$$= e^{-ax^2} e^{bx} \quad \text{s.t. } a \geq 0,\ b \in \mathbb{R}$$

do this later



$E(X) = \theta$

$Var(X) = \sigma^2$

$Supp(X) = \mathbb{R}$

Param space is two dimensional.

$\theta \in \mathbb{R}$

$\sigma^2 \in (0, \infty)$

given $X_1, \ldots, X_n \overset{iid}{\sim} N(\theta, \sigma^2)$ and

Goal: pretend σ^2 is known, infer θ. First find MLE

$$\mathcal{L}(\theta; x, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x_i - \theta)^2}$$

$$= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{1}{2\sigma^2}\sum(x_i - \theta)^2}$$

$(x_i - \theta)^2 = x_i^2 - 2\theta x + \theta^2$

$\Rightarrow \sum(x_i - \theta)^2 = \sum x_i^2 - 2\theta \sum x + \sum \theta^2$

$= \sum x_i^2 - 2\theta \bar{x} n + n\theta^2$

$$= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{1}{2\sigma^2}(\sum x_i^2 - 2\theta \bar{x} n + n\theta^2)}$$

$$= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{\sum x_i^2}{2\sigma^2}} e^{\frac{\theta \bar{x} n}{\sigma^2}} e^{-\frac{n\theta^2}{2\sigma^2}}$$

$$\ell(\theta; x, \sigma^2) = n \ln\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) - \frac{\sum x_i^2}{2\sigma^2} + \frac{\theta \bar{x} n}{\sigma^2} - \frac{n\theta^2}{2\sigma^2}$$

$$\ell'(\theta; x, \sigma^2) = \frac{\bar{x} n}{\sigma^2} - \frac{2n\theta}{2\sigma^2} \overset{set}{=} 0$$

$$\Rightarrow \bar{x} n - n\theta = 0$$

$$\Rightarrow \bar{x} - \theta = 0$$

$$\Rightarrow \boxed{\hat{\theta}_{mle} = \bar{x}}$$

Fun with kernels...

Just one..

$$P(X|\theta,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x_i^2}{2\sigma^2}} e^{\frac{\theta x_i}{\sigma^2}} e^{-\frac{\theta^2}{2\sigma^2}}$$

$$\propto \underbrace{e^{-\frac{x_i^2}{2\sigma^2}} e^{\frac{\theta x_i}{\sigma^2}}}_{e^{-a x_i^2} e^{bx}}$$

s.t. $a \geq 0$, $b \in \mathbb{R}$

$$P(X_{whole}|\theta,\sigma^2) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{\Sigma x_i^2}{2\sigma^2}} e^{\frac{\theta \bar{x} n}{\sigma^2}} e^{-\frac{n\theta^2}{2\sigma^2}}$$

$$\propto e^{-\frac{\Sigma x_i^2}{2\sigma^2}} e^{\frac{\theta \bar{x} n}{\sigma^2}}$$

$$= e^{-a\Sigma x_i^2} e^{b\bar{x}} \quad \text{s.t. } a \geq 0,\ b \in \mathbb{R}$$

Now... let's figure out what the prior could be... back to Bayes Rule

$$P(\theta|X,\sigma^2) = \frac{P(X|\theta,\sigma^2)\, P(\theta|\sigma^2)}{P(X|\sigma^2)}$$

Why? σ^2 known.. so it's conditional on everything...

$$\propto P(X|\theta,\sigma^2)\, P(\theta|\sigma^2)$$

$$= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{\Sigma x_i^2}{2\sigma^2}} e^{\frac{\theta \bar{x} n}{\sigma^2}} e^{-\frac{n\theta^2}{2\sigma^2}} P(\theta|\sigma^2)$$

$$\propto \underbrace{e^{\frac{\theta \bar{x} n}{\sigma^2}} e^{-\frac{n\theta^2}{2\sigma^2}}}_{e^{-a\theta^2} e^{b\theta}} P(\theta|\sigma^2)$$

s.t. $a = \frac{n}{2\sigma^2}$, $b = \frac{\bar{x} n}{\sigma^2}$

What should $P(\theta|\sigma^2)$ be?

Let's match the kernel like we've been doing...

The kernel is a normal so let's do a normal...

$$P(\theta|\sigma^2) = N(\mu_0, \tau^2) = \frac{1}{\sqrt{2\pi\tau^2}} e^{-\frac{1}{2\tau^2}(\theta-\mu_0)^2}$$

$$\propto e^{-\frac{1}{2\tau^2}(\theta^2 - 2\theta\mu_0 + \mu_0^2)}$$

$$\propto e^{-\frac{\theta^2}{2\tau^2}} e^{\frac{\theta\mu_0}{\tau^2}}$$

$$\Rightarrow P(\theta|X,\sigma^2) \propto e^{-\frac{n}{2\sigma^2}\theta^2} e^{\frac{\bar{X}n}{\sigma^2}\theta} e^{-\frac{1}{2\tau^2}\theta^2} e^{\frac{\mu_0}{\tau^2}\theta}$$

$$= e^{-\left(\frac{n}{2\sigma^2}+\frac{1}{2\tau^2}\right)\theta^2} e^{\left(\frac{\bar{X}n}{\sigma^2}+\frac{\mu_0}{\tau^2}\right)\theta}$$

This is normal distr since the kernel is normal

but which one is it?

$$\propto \frac{1}{\sqrt{2\pi v^2}} e^{-\frac{1}{2v^2}(\theta-c)^2}$$

$$\propto e^{-\frac{1}{2v^2}(\theta-c)^2} = e^{-\frac{1}{2v^2}(\theta^2-2c\theta+c^2)} = e^{-\frac{1}{2v^2}\theta^2} e^{\frac{c}{v^2}\theta} e^{-\frac{c^2}{2v^2}}$$

$$\propto e^{-\frac{1}{2v^2}\theta^2} e^{\frac{c}{v^2}\theta}$$

solve for v^2, c

$$\Rightarrow \frac{1}{\cancel{2}v^2} = \frac{n}{\cancel{2}\sigma^2} + \frac{1}{\cancel{2}\tau^2} \Rightarrow v^2 = \frac{1}{\frac{n}{\sigma^2}+\frac{1}{\tau^2}}$$

$$\Rightarrow \frac{c}{v^2} = \frac{\bar{X}n}{\sigma^2} + \frac{\mu_0}{\tau^2}$$

$$\Rightarrow c = \left(\frac{\bar{X}n}{\sigma^2}+\frac{\mu_0}{\tau^2}\right)v^2 = \frac{\frac{\bar{X}n}{\sigma^2}+\frac{\mu_0}{\tau^2}}{\frac{n}{\sigma^2}+\frac{1}{\tau^2}}$$

$$\propto N\left(\underbrace{\frac{\frac{\bar{X}n}{\sigma^2}+\frac{\mu_0}{\tau^2}}{\frac{n}{\sigma^2}+\frac{1}{\tau^2}}}_{\theta'}, \underbrace{\frac{1}{\frac{n}{\sigma^2}+\frac{1}{\tau^2}}}_{\sigma^{2\prime}}\right)$$

Normal is conjugate prior for normal likelihood family.