⊙ Theoretical Q's about the exponential family of distribution

MODULE 1

- use i index to specify that each observation might have different parameters and weights, but suppress <u>here</u> for simplicity (since likelihood <u>not</u> focus).

$$f(y) = \exp\left\{ \frac{y\theta - b(\theta)}{\phi} \cdot w + c(y, \phi, w) \right\}$$

1) Binomial distribution

- Bernoulli process:
  - a) experiment made up of independent trials
  - b) each trial: success or failure ("A")
  - c) $P(A)$ is the same in each trial

If we perform n trials the $Y = \#$ successes is binomial (n, p).

- $f(y) = \binom{n}{y} p^y (1-p)^{n-y}$

$$= \exp\left(\ln\binom{n}{y} + y \ln p + (n-y)\ln(1-p) \right)$$

$$= \exp\left(y\,[\ln p - \ln(1-p)] + n\ln(1-p) + \ln\binom{n}{y} \right)$$

$$= \exp\Big\{ y \cdot \underbrace{\ln\left(\tfrac{p}{1-p}\right)}_{\theta} + \underbrace{n\ln(1-p)}_{-b(\theta)} + \underbrace{\ln\binom{n}{y}}_{c(y,\phi,w)} \Big\}$$

* $\theta = \ln\left(\frac{p}{1-p}\right)$

then for $b(\theta)$

$$e^{\theta} = \frac{p}{1-p} \Leftrightarrow (1-p)e^{\theta} = p$$

$$e^{\theta} - pe^{\theta} = p$$

$$e^{\theta} = p + pe^{\theta} = p(1+e^{\theta})$$

$$\underline{p = \frac{e^{\theta}}{1+e^{\theta}}}$$

so we have $\underline{(1-p)} \Rightarrow 1-p = 1 - \frac{e^{\theta}}{1+e^{\theta}} = \frac{1+e^{\theta}-e^{\theta}}{1+e^{\theta}} = \underline{\frac{1}{1+e^{\theta}}}$

and $\underline{b(\theta)} = -n\ln(1-p) = -n\ln\left(\frac{1}{1+e^{\theta}}\right) =$

$$-n\underbrace{\ln 1}_{0} + n\ln(1+e^{\theta}) = \underline{n\ln(1+e^{\theta})}$$

Then $\underline{\phi = 1}$ and $\underline{w = 1}$ and $\underline{c(y,\phi,w) = \ln\binom{n}{y}}$

Textbook: Bernoulli $(n=1)$ has

$\theta = \ln\frac{p}{1-p}$ $\quad b(\theta) = \ln(1+e^{\theta})$ and $\phi = 1$

Missing, the E and Var

$E(Y) = \underline{n \cdot p}$ and $Var(Y) = \underline{n \cdot p(1-p)}$ [known] and

$b(\theta) = n \cdot \ln(1+e^{\theta})$ where $\theta = \log \frac{p}{1-p} = \text{logit}(p)$

$$b'(\theta) = \frac{db}{d\theta} = \frac{n}{1+e^{\theta}} \cdot e^{\theta} = \frac{n \cdot e^{\theta}}{1+e^{\theta}} = \underline{n \cdot p}$$

since $p = \frac{e^{\theta}}{1+e^{\theta}}$

$$b''(\theta) = \frac{d}{d\theta} b'(\theta) = \frac{ne^{\theta} \cdot (1+e^{\theta}) - ne^{\theta} \cdot e^{\theta}}{(1+e^{\theta})^2}$$

$$= \underbrace{\frac{ne^{\theta}}{1+e^{\theta}}}_{p} \cdot \underbrace{\frac{1+e^{\theta} - e^{\theta}}{(1+e^{\theta})}}_{\frac{1}{1+e^{\theta}} = 1-p} = \underline{np(1-p)}$$

$\Uparrow$ since

$b''(\theta) \cdot \frac{\phi}{w}$ $\Leftarrow$ $\Uparrow$ 1

"

$Var(Y)$

2) Poisson distribution

a) Poisson process: we observe events occuring within a time interval or area in space.

- the number of events occuring in one interval or aree is independent of the number of events in disjoint intervals/areas
- the probability of an event inside an interval or area is proportional to the length of the interval or size of the area.
- the probability that more then one event occur in a small interval or area is negligible.

$\Rightarrow$ we have a Poisson process.

Then: Y = "the number of events within an intervall or area" follows a Poisson distribution.

$$f(y) = \frac{\mu^y}{y!} e^{-\mu}, \text{ where } E(Y) = Var(Y) = \mu$$

$$f(y) = \frac{\mu^y}{y!} e^{-\mu} = \exp\left(-\mu + y\ln\mu - \ln y!\right)$$

$$= \exp\left(y \cdot \underbrace{\ln\mu}_{\theta} - \underbrace{\mu}_{b(\theta)} + \underbrace{(-\ln y!)}_{c(y)} \right)$$

so exp. fam with $\theta = \ln\mu \iff \mu = \exp\theta$

$b(\theta) = \exp\theta$

$c(y) = -\ln y!$

$$E(Y) = \frac{db}{d\theta} = \frac{d}{d\theta}\exp\theta = \exp\theta = \mu$$

$$Var(Y) = \frac{d^2b}{d\theta^2} = \frac{d}{d\theta}\exp\theta = \exp\theta = \mu$$

3) Normal distribution

Measurements of physical properties or scientific measurement (with error) is known to be normally distributed. This migth be thought of as a version of the central limit theorem.

$$f(y) = \frac{1}{\sqrt{2\pi}\,\sigma} \cdot \exp\left\{-\frac{1}{2}\frac{(y-\mu)}{\sigma^2}\right\}$$

and $E(Y) = \mu$, $Var(Y) = \sigma^2$

$$f(y) = \exp\left\{\left(-\frac{1}{2}\sigma^{-2}\right)\left(y^2 - 2\mu y + \mu^2\right) - \ln\left(\sqrt{2\pi}\,\sigma\right)\right\}$$

$$= \exp\left\{\left(-\frac{1}{2\sigma}\right)\left(-2\mu \cdot y + \mu^2\right) + \left(-\frac{1}{2\sigma}\right)y^2 - \ln\left(\sqrt{2\pi}\,\sigma\right)\right\}$$

$$= \exp\left\{\underbrace{\frac{1}{\sigma^2}}_{\frac{1}{\phi}} \cdot \underbrace{\mu \cdot y}_{\theta \cdot y} - \underbrace{\frac{1}{\sigma^2}}_{\frac{1}{\phi}} \cdot \underbrace{\frac{1}{2}\mu^2}_{b(\theta)} + \underbrace{\left(-\frac{1}{2}\frac{1}{\sigma^2}y^2 - \ln\sqrt{2\pi}\,\sigma\right)}_{c(y,\phi)}\right\}$$

$\theta = \mu$, $\phi = \sigma^2$, $w = 1$, $c(y, \phi)$

$\Rightarrow$ exponential family $\qquad b(\theta) = \frac{1}{2}\theta^2$

$$E(Y) = \frac{d}{d\theta} b(\theta) = \frac{d}{d\theta}\left(\frac{1}{2}\theta^2\right) = \theta = \underline{\underline{\mu}}$$

$$\frac{d^2 b(\theta)}{d\theta^2} = \frac{d}{d\theta}\theta = 1$$

$$V_\theta(Y) = \frac{d^2 b(\theta)}{d\theta^2} \cdot \frac{\phi}{w} = 1 \cdot \frac{\sigma^2}{1} = \underline{\underline{\sigma^2}}$$

4) Gamma distribution

The waiting time until the νth event in a Poisson process has a Gamma (or Erlang) distribution. When $\nu=1$ the exponential distribution occurs - and is much used in survival analysis. The χ^2_δ-distribution is a special case of the gamma distribution ($\frac{\nu}{\mu}=2$ and $\nu=\frac{\delta}{2}$ with δ as param. in the χ^2 - but usually ν, so confusing).

Use parameterization on page 643 of appendix B

$$f(y)=\frac{1}{\Gamma(\nu)}\left(\frac{\nu}{\mu}\right)^{\nu}y^{\nu-1}\exp\left(-\frac{\nu}{\mu}y\right),\quad y>0$$

$$f(y)=\exp\left\{-\frac{\nu}{\mu}\cdot y+\nu\ln\left(\frac{\nu}{\mu}\right)+(\nu-1)\cdot\ln y-\ln(\Gamma(\nu))\right\}$$

$$=\exp\left\{\frac{-\frac{1}{\mu}\cdot y}{\frac{1}{\nu}}+\frac{\ln\nu+\ln\left(\frac{1}{\mu}\right)}{\frac{1}{\nu}}+(\nu-1)\ln y-\ln(\Gamma(\nu))\right\}$$

$$=\exp\left\{\frac{\underbrace{-\frac{1}{\mu}}_{\theta}\cdot y+\ln\left(--\frac{1}{\mu}\right)}{\underbrace{\frac{1}{\nu}}_{\phi}}+\overbrace{\nu\cdot\ln\nu+(\nu-1)\ln y-\ln(\Gamma(\nu))}^{c(y,\phi)}\right\}$$

$w=1$

$b(\theta)=-\ln\left(--\frac{1}{\mu}\right)=-\ln(-\theta)$

$\theta=-\frac{1}{\mu}\Leftrightarrow\mu=-\frac{1}{\theta}$

$$E(Y) = b'(\theta) = \frac{d}{d\theta}\left(-\ln(-\theta)\right) = -\frac{1}{-\theta}(\cdot -1) = \underline{\underline{-\frac{1}{\theta} = \mu}}$$

$$\frac{d^2 b(\theta)}{d\theta^2} = \frac{d}{d\theta}\left(-\frac{1}{\theta}\right) = \frac{1}{\theta^2} = \mu^2$$

$$Var(Y) = \frac{d^2 b(\theta)}{d\theta^2} \cdot \frac{\phi}{w} = \frac{1}{\theta^2} \cdot \frac{\frac{1}{\nu}}{1} = \underline{\underline{\frac{1}{\nu\theta^2} = \frac{\mu^2}{\nu}}}$$