Reminder · 1st Midterm Oct. 9

- aids allowed · calculator
  - 8½" × 11" two-sided formula sheet

Review of last day

Poisson Regression Model
- log-linear model

Y - Poisson $\mu = E(Y)$
· relate to $X_1 \ldots X_p$

$* \log(\mu) = \eta = (\text{offset}) + \beta_0 + \beta_1 X_1 + \cdots \beta_p X_p$
$-\infty < \quad < +\infty$

→ log → example link function in generalized linear model

also · logistic regression
link = logit = $\log \frac{p}{1-p}$

estimation max. likelihood

$- X^t(\underset{\sim}{Y} - \underset{\sim}{\mu}) = 0$ → likelihood equation

$\underset{\sim}{\hat{\beta}}$ → asymptotically unbiased for $\underset{\sim}{\beta}$

asymptotic · variance · covariance
- inverse Fisher Information

following Simonoff's ~~data~~ notation

W = diagonal matrix · $n \times n$

with $\hat{\mu}$ on diagonal

asy var. cov. of $\hat{\beta} = (X^t W X)^{-1}$

→ plug in $\hat{\beta}$ → yields s.e.'s.

ex. Shark attack

β_1 = "slope" by year

$$\log\left(\frac{\mu}{\underbrace{\text{offset}}_{\text{pop}^n \text{ denominator}}}\right) = \beta_0 + \beta_1 Yr$$

λ = rate

$$\underset{\text{rate}}{\lambda} = e^{\beta_0} \times (e^{\beta_1})^{Yr}$$

e^{β_1} = yearly proportional increase

- Relative Risk for one year ~~lapse~~ progression

To test $\beta_1 = 0$

→ no change year to year

R

Confidence Intervals

"Confint"

```
Estimates with confidence intervals
-----------------------------------
               Estimate  Std. Error     Pr(>|z|)        2.5 %       97.5 %
(Intercept) -75.79226904 8.658220532 2.062996e-18 -92.76206945 -58.82246862
year          0.03117395 0.004361182 8.801855e-13   0.02262619   0.03972171

Back-transformed estimates with confidence intervals
----------------------------------------------------
               Estimate  Std. Error     Pr(>|z|)        2.5 %       97.5 %
(Intercept) 1.212930e-33 8.658220532 2.062996e-18 5.175414e-41 2.842670e-26
year        1.031665e+00 0.004361182 8.801855e-13 1.022884e+00 1.040521e+00

Back-transformed estimates with profile confidence Intervals
------------------------------------------------------------
               Estimate  Std. Error     Pr(>|z|)        2.5 %       97.5 %
(Intercept) 1.212930e-33 8.658220532 2.062996e-18 3.775357e-41 2.113763e-26
year        1.031665e+00 0.004361182 8.801855e-13 1.023034e+00 1.040684e+00
```

log scale

exp.

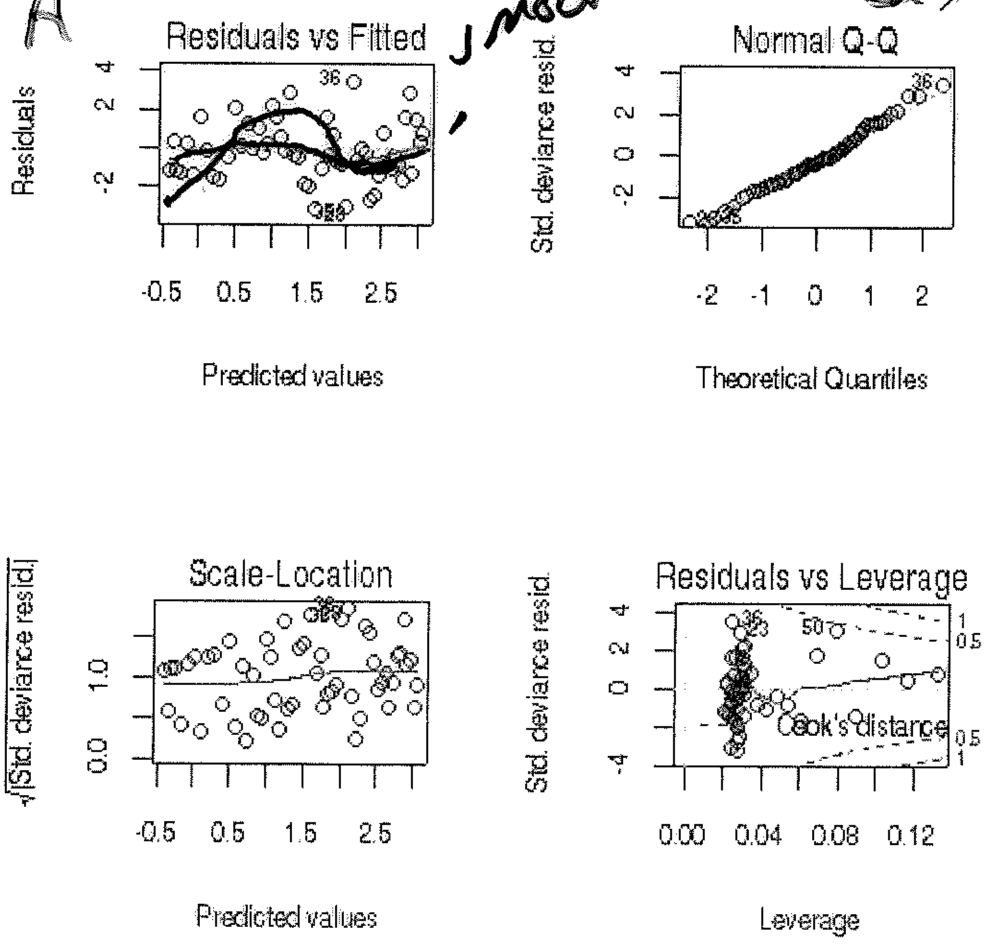
$e^{\hat{\beta}}$

Examining Assumptions

– See web page for function

smoother (in red)

A



Test $H_0: \beta_1 = 0$ $\quad z = \frac{\hat{\beta}}{s.e.} = \frac{.0312}{.0044} = 7.1$

$p < .0001$

For parameter interpretat

→ need to $e^{\hat{\beta}_1}$

→ need to focus on conf. intervals

endpoints of conf. interval for e^{β_1} is exp(endpoints for β_1)

s.e.(~~$\hat{\beta}_1$~~ $e^{\hat{\beta}_1}$)

→ use Delta method

$$s.e.(e^{\hat{\beta}_1}) = s.e.(\hat{\beta}_1) \times \underbrace{\frac{\partial e^{\beta_1}}{\partial \beta_1}\Bigg|_{\beta_1 = \hat{\beta}}}_{e^{\hat{\beta}_1}}$$

- could use this in conf. int.
- theoretically but better to transform "raw" conf. interval

earlier

$e^{\hat{\beta}_1} = 1.\underbrace{032}$: rate year i+1 $\approx 1.032 \times$ rate year i

- roughly 3% increase per year

Conf. int. 2% to 4%

note ~~$e^{(1+}$~~ $e^{\delta} \approx 1 + \delta$ for δ small

→ if $\hat{\beta}_j$ is small, can interpret as. $100 \times \hat{\beta}_j$ as estimated % rate of increase per unit of X_j

~~Next~~ Note on confint function in R

→ generic fla for C.I

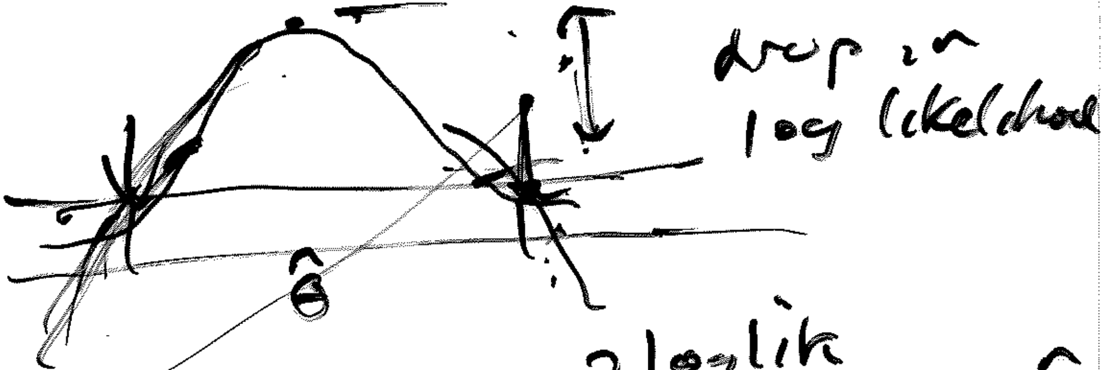
$\hat{\Theta} \pm z_{\alpha/2}\,\mathrm{s.e.}(\hat{\theta})$ Wald interval

Note. 3 generic likelihood based approaches

1: Wald - asympt. Var of $\hat{\beta}$

2 log likelihood approach

→ likelihood ratio tests



-2 loglik $\hat{}$ χ^2 dist

- chose this drop to correspond to signif at 5%

→ profile likelihood approach

- thought to be superior to Wald

→ default in confint.

All inferences depend on assumptions of model

- log-linearity . $\log(\mu) = \underset{\text{sometimes } 0}{\underset{\downarrow}{\text{offset}}} + \beta_0 + \beta_1 X_1 \ldots$
- variance $Var(y_i) = \mu_i$
- independence → depends mainly on data collection process
- y_i ~ Poisson - not essential in large samples

will base examination on
residuals $: y_i - \hat{y}_i$
or $y_i - \hat{\mu}_i$

~~in~~ to allow for non-homogeneous variance
we scale residuals by
$\sqrt{\mathrm{Var}(y_i)}$; estimated by $\sqrt{\hat{\mu}_i}$

Pearson residual $r_i^P = \frac{y_i - \hat{\mu}_i}{\sqrt{\hat{\mu}_i}}$

alternative (less intuitive)
approach to defining residual
motivated by considering log
likelihood

in Normal -2 log likelihood
case . evaluated at $\hat{\beta}$

$$= \frac{\sum (y_i - \hat{y}_i)^2}{\sigma^2} = \frac{SS(Resid)}{\sigma^2}$$

so this suggests trying to
express -2 log lik $= \sum (r_i)^2$
(evaluated at $\hat{\beta}$)

Last day we saw

$-2 \log lik : -2$ ~~$\sum y_i e^{\hat{\mu}_i}$~~

$$= -2 \sum y_i \log(\hat{\mu}_i) - \hat{\mu}_i - \boxed{\log(y_i!)}$$

To "simplify" thing . we need to standardize -2 log lik to get rid of arbitrary "constant" terms.

Introduce a standard benchmark model - saturated model

- perfect fit . $\hat{\mu}_i = y_i$
  - equivalent to a model with 1 β / obsn

Deviance for a fitted model

$$= -2 \log lik(\hat{\beta}) - 2 \log lik$$

for ~~st~~ saturated model

Poisson deviance fla

$$= -2 \sum \underbrace{y_i \log\left(\frac{y_i}{\hat{\mu}_i}\right) - (y_i - \hat{\mu}_i)}_{d_i}$$

$= -2 \sum d_i$ - d_i = deviance component

: $\approx (y_i - \hat{y}_i)^2$

deviance residual

: $r_i^d = \sqrt{d_i} \, \text{sign}(y_i - \hat{\mu}_i)$

→ current favourite

Using residuals

- in R llfit ← glm(......)

r ← residual(llfit)

- various options - type
- default is "deviance"

used in plot.(llfit)

- yields 4 plots as default

A - r_i^d vs $\hat{\eta}_i = \log(\hat{\mu}_i)$

$\downarrow$

linear predictor

(linear in β's)

- examine linearity overall
- If only 1 x → quite usef

if > 1 Plot r_i^d vs X_j's separately

- perhaps indicator of non-linearity