Stat 534: formulae referenced in lecture, week 4:
Profile likelihood, Lincoln-Petersen estimator, extension to more than 2 sampling occasions

Profile likelihood: another way to compute a confidence interval

- The asymptotic confidence interval, $\hat{\theta} \pm z_{1-\alpha/2} \sqrt{\text{Var } \hat{\theta}}$, assumes that $\hat{\theta} \sim N()$
  - Theory $\Rightarrow$ the distribution of a mle, $\hat{\theta}$, converges to a normal distribution as the sample size increases
  - Doesn't $\Rightarrow$ the distribution is normal when # observations is small-moderate.
- Profile likelihood intervals do not rely on asymptotic normality

Relationship between a hypothesis test and a confidence interval

- A useful applied statistics reminder: holds for any test and associated interval
- Consider a 1 parameter model, e.g. Y Poisson (λ)
  - have mle, $\hat{\lambda}$, e.g., 16.4 for the moist woodland oaks
  - want to test $H_o : \lambda_0 = 15$
  - Likelihood ratio test statistic is $C = -2\left[\text{lnL}(\lambda_0) - \text{lnL}(\hat{\lambda})\right]$
  - Reject H_o, $\lambda_0 = 15$ at $\alpha = 0.05$ when $C > 3.84$
  - lnL(15) = 1.27, so accept H_o, $p > 0.05$ (p = 0.52, if you want to know)
  - What about other null hypothesis values?

λ_0	lnL	C	
16.4	-3.47	0	accept H_o
15.0	-4.10	1.27	accept H_o
14.5	-4.66	2.39	accept H_o
18.5	-4.71	2.48	accept H_o
19.0	-5.34	3.73	accept H_o
14.0	-5.42	3.90	reject H_o, $p < 0.05$
19.5	-6.08	5.21	reject H_o, $p < 0.05$

  - illustrated in profile1.pdf
  - 95% confidence interval for λ is all the λ_0 values for which test $\Rightarrow p > 0.05$
  - Need to find the λ_0 values where $p = 0.05$. Those are the end points of the 95% ci.

λ_0	lnL	C
14.015	-5.39	3.84
19.04	-5.39	3.84

What about models with 2 or more parameters?

- Still want a ci for one parameter
  - Aside: generalizes to confidence regions for two or more parameters simultaneously
- Example: redpoll data, 2 sampling occasions, same capture probability each time, p
- illustrated in profile2.pdf: lnL as function of (N, p) for redpoll data
- mle's are $\hat{N} = 90.5$, $\hat{p} = 0.32$
- want ci for N
- Again, start with a test
  - E.g., test $H_o : N_0 = 80$
  - This null hypothesis says nothing about p
    need to compare lnL($N = 90.5$, $p = 0.32$) to lnL($N = 80$, $p =$??).
  - Need to find the mle for p given $N_0 = 80$.
  - Sometimes have equation for this, usually have to numerically optimize
  - Test statistic is $C = -2 \left[\text{lnL}(N_0, \hat{p} \mid N_0) - \text{lnL}(\hat{N}, \hat{p}\right]$
  - Null hypothesis has 1 "free" parameter, alternative has 2, so C has 1 df
  - 95% ci: Find the set of N_0 values where $p > 0.05$, i.e. $C < 3.84/2$

Connections between profile likelihood intervals and asymptotic normality (Wald) intervals

- Everything below is for 95% ci = $\alpha = 5\%$ test, generalizes to other %'s
- Wald interval is based on a Z-test: reject $H_0 : N = N_0$ when $Z = \frac{\hat{N} - N_0}{\sqrt{\text{Var } \hat{N}}} > 1.96$
  - Assumes that $\hat{N}$ has a normal distribution
  - Equivalent to a profile likelihood plot that is quadratic around $\hat{N}$
- Profile likelihood interval is based on a likelihood ratio test (LRT)
  - Does not assume a distribution for $\hat{N}$
  - The Chi-square distribution for C in a LRT is an asymptotic result
  - But practical experience + theory $\Rightarrow$ the Chi-square assumption is appropriate for much smaller sample sizes than the normality assumption

What if we used $\log N$ as the parameter, not N?

- lnL is unchanged.

- Invariance property of mle's:
  - lnL unchanged for a general transformation of parameters, not just log
- Profile likelihood plot usually looks different
- But, profile likelihood confidence intervals are identical
- Asymptotic normal intervals are not

Importance of number of recaptures

- Define $\mu = \frac{n_1 n_2}{N} = \mathrm{E}\ (m2 \mid n1, n2)$
- c.v. of $\hat{N} \approx 1/\sqrt{\mu}$
- plug-in estimate is $1/\sqrt{m_2}$

More than two sampling occasions:

- Improves precision
- Allows considering multiple models for capture process
- Consider 3 occasions. Data can be summarized as $2^3 = 8$ possible capture histories
- Do not observe the NNN history.
- Goal will be to estimate n_{000}

Time			
1	2	3	# animals
Y	Y	Y	n_{111}
Y	Y	N	n_{110}
Y	N	Y	n_{101}
Y	N	N	n_{100}
N	Y	Y	n_{011}
N	Y	N	n_{010}
N	N	Y	n_{001}
N	N	N	n_{000}

- Extends in obvious way to more than two capture occasions
  - 2^k possible histories for k occasions
  - Don't know # unseen, $n_{0\cdots0}$. All other counts are observed.

Otis models for capture process in a closed population:

- Named by Dave Otis in his PhD thesis. Published in a wildlife monograph.
  - Dave was ISU Coop unit leader before Bob Klaver, late 1990's? to mid-2010's?
- Three issues that could be included in a model for capture probability
  - All concern p, the probability that an individual is captured on an occasion
  - T: Variability between occasions (times)
  - B: Behavioural variability
  - H: Individual heterogeneity
- Individually or in combination
- Models named by subscripts
  - M0: same capture probability for all sampling occasions
  - Mt: capture probability differs between occasions
  - Mb: behavioural response: capture probability differs between 1st capture and subsequent ones (trap happy / trap shy behaviour)
  - Mh: individuals have different capture probabilities
  - and all combinations, Mtb, Mth, Mbh, Mtbh
  - Heterogeneity between individuals (model names with h) makes life very difficult, will discuss last

M0: constant capture probability

- lnL depends on 3 summary statistics
  - t: # sampling occasions
  - $n_.$: total number of captures $= \sum_{i=1}^{t} n_i$
  - M_{t+1} : total number of marked individuals after the last sampling occasion
    = total number of animals seen at least once (perhaps more often)
- the important part of the log likelihood depends on these quantities (the sufficient statistics), not any additional details, e.g., individual n_i

$$\text{lnL} = \log N! - \log \text{constant} - \log(N - M_{t+1})! + n_. \log p + t\, N - n_. \log(1-p)$$

- constant depends on the data, not on any parameter