2 Profile Likelihood

The term "profile likelihood" can mean multiple things.

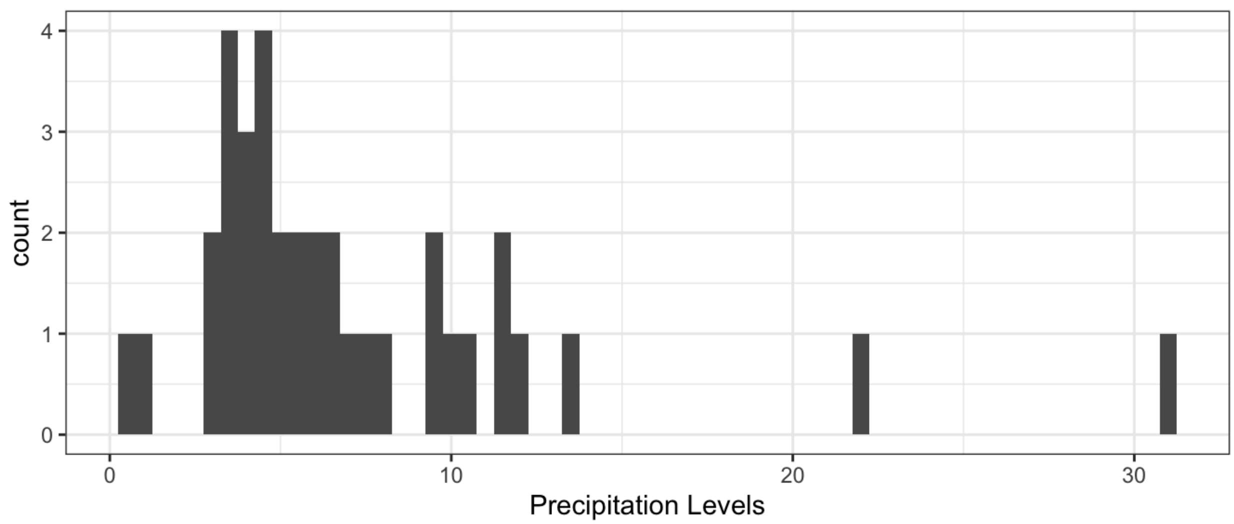
In all cases, this is a way to get a point estimate and assess uncertainty in a proportion of the parameter vector while essentially ignoring the other parameters.

2.1 Analytical Methods via Profile Likelihoods

In certain problems it is possible to maximize the log likelihood for part of $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^\top, \boldsymbol{\theta}_2^\top)^\top$ without actually knowing the value of the other part.

The **profile likelihood** is the usual likelihood with the known function of part of the parameter vector inserted for that parameter, making the likelihood only a function of one part of the vector.

Example (Hurricane Data, Cont'd): For 36 hurricanes that had moved far inland on the East Coast of the US in 1900-1969, maximum 24-hour precipitation levels during the time they were over mountains.



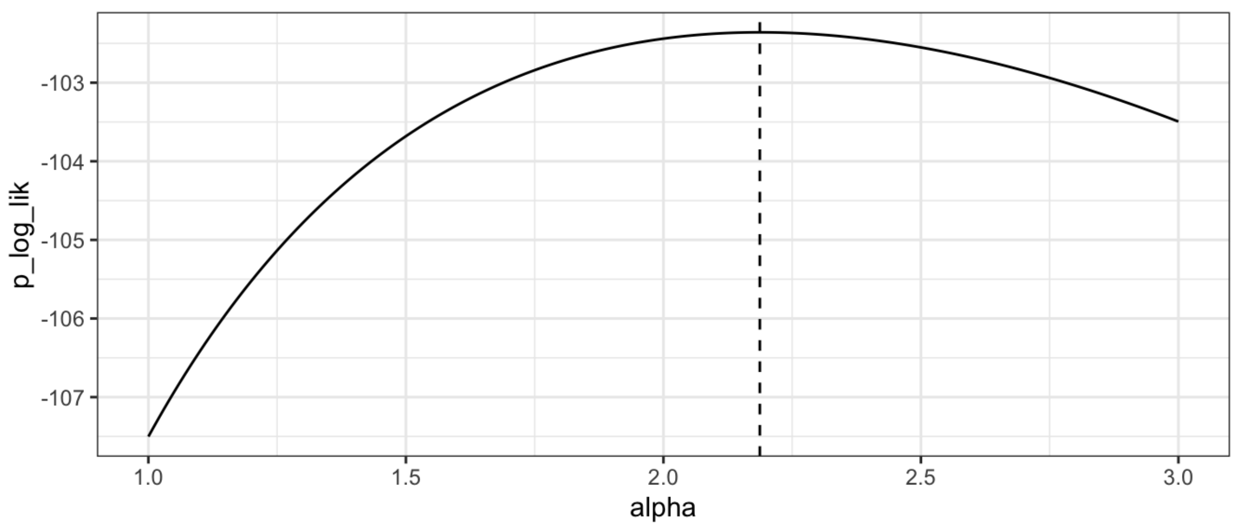
We modeled the precipitation levels with a gamma distribution, which has log likelihood

$$\ell(\alpha, \beta) = -n \log \Gamma(\alpha) - n\alpha \log \beta + (\alpha - 1) \sum \log Y_i - \frac{\sum Y_i}{\beta}$$

```r
gamma_prof_loglik <- function(alpha, data) {
  beta <- mean(data) / alpha
  sum(dgamma(data, alpha, scale = beta, log = TRUE))
}

## get maximum profile likelihood estimate
alpha_mple <- optim(1, gamma_prof_loglik, data = hurr_rain, method =
        "BFGS", control = list(fnscale = -1))

## plot profile likelihood
data.frame(alpha = seq(1, 3, length.out = 200)) |>
  rowwise() |>
  mutate(p_log_lik = gamma_prof_loglik(alpha, hurr_rain)) |>
  ggplot() +
  geom_line(aes(alpha, p_log_lik)) +
  geom_vline(aes(xintercept = alpha_mple$par), lty = 2)
```



2.2 Numerical Methods via Profile Likelihoods

The log likelihood can be maximized over one portion of the partition $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^\top, \boldsymbol{\theta}_2^\top)^\top$ for any fixed value of the other, even if that maximization cannot be expressed as an explicit function.

We can define a profile likelihood as

$$L^p(\boldsymbol{\theta}_2) = \max_{\boldsymbol{\theta}_1} L(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2).$$

The profile likelihood and log profile likelihood behave in many ways like true likelihood functions:

1. The estimate of $\boldsymbol{\theta}_2$ found by maximizing $L^p(\boldsymbol{\theta}_2)$ is the MLE of $\boldsymbol{\theta}_2$.

2. A likelihood ratio test statistics formed with the profile likelihood has a limiting χ^2 distribution.

3. A profile likelihood confidence region is a valid approximate confidence region for $\boldsymbol{\theta}_2$.

Where does this confidence region come from?

However, these are *not* full likelihood functions.