Maximum likelihood Lab 3: Stock-Recruitment

answer key

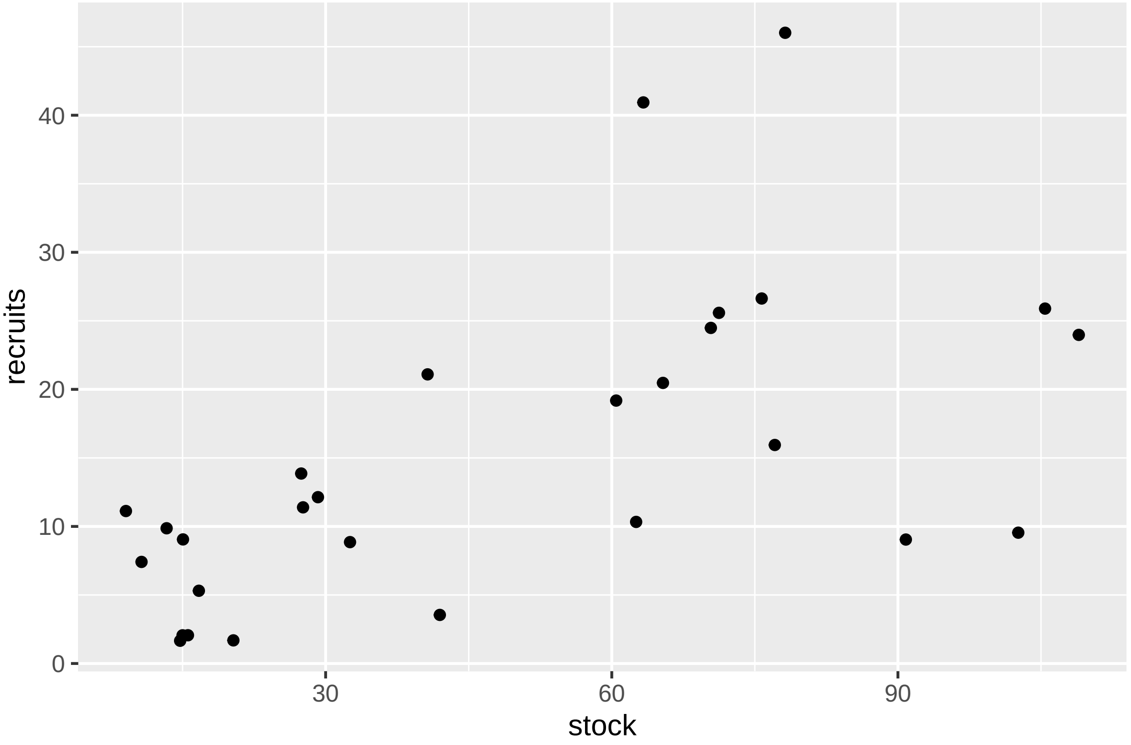
BIOHOPK 143H - Winter 2021

The Data

We'll be using data on lake trout stock and recruitment at Gull Island Shoal, lake superior, from 1964 - 1991. In this data, the `stock` column is a geometric mean of the number of adult female lake trout from surveys, and the `recruits` column is a geometric mean of the number of recruits (age 0 trout) from surveys.

When we load and plot the data, we can see that there hardly looks to be a relationship, as is fairly typical of stock-recruitment data - there's more noise than "signal."

```
LakeTroutGIS %>%
  ggplot(aes(stock, recruits)) +
  geom_point()
```



Stock-Recruitment Models

Density-Independence

In this lab, we're going to fit three different stock-recruitment relationships to these data. The first is a model assuming **density independence**, where the number of recruits per spawner is constant regardless of the size of the spawning stock, and there are no effects on recruitment from crowding or competition. This model is:

$$R = aS$$

Where R is the number of recruits, S is the number of adult female fish, and a is the number of recruits arising from each female adult, which of course is proportional to the fecundity.

Beverton-Holt

As you know, there are multiple ways to introduce density-dependence into this model. One option is the **Beverton-Holt** model:

$$R = \frac{aS}{1 + bS}$$

In this model, a describes the number of recruits per spawner at low biomass ($S \approx 0$), and b is a parameter that controls the strength of density dependence - if there is *no* density dependence, than $b = 0$, and this model collapses to the density independent model. The peak recruitment under this model occurs at the asymptote, $R_p = a/b$.

Ricker

While in the Beverton-Holt model recruitment reaches an asymptote as the spawning stock size increases, the Ricker model allows for recruitment to decline at high stock biomass. The stock-recruitment relationship for the Ricker model is:

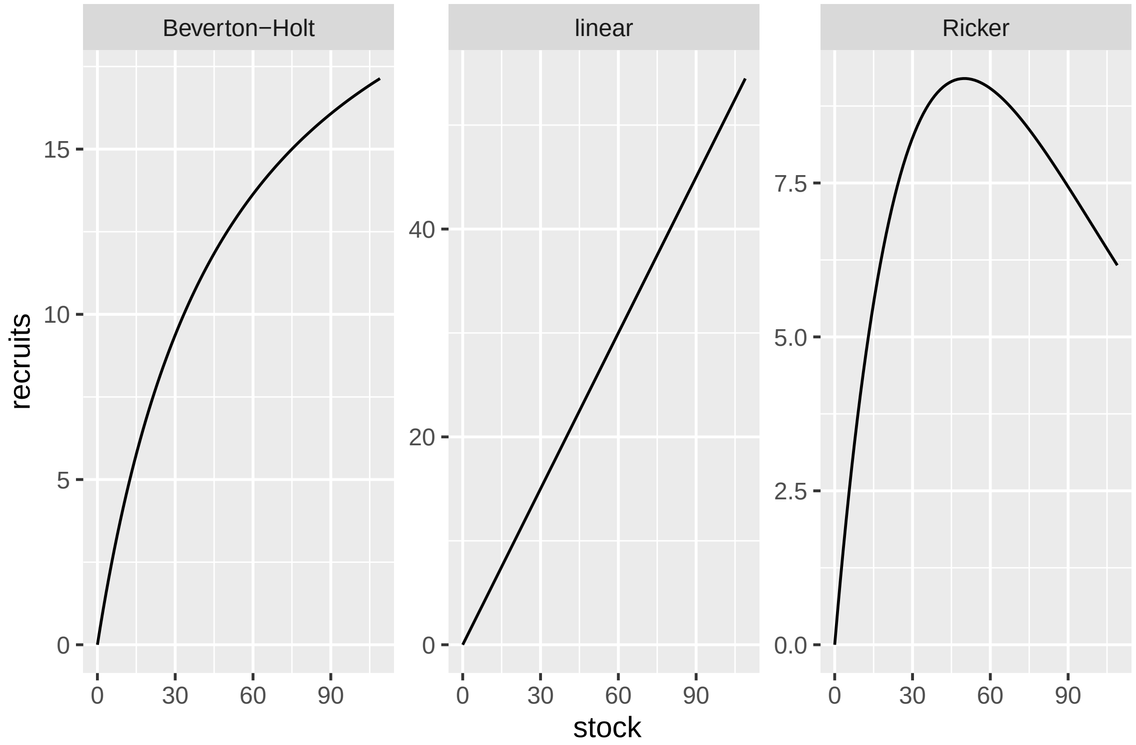
$$R = aSe^{-bS}$$

Where, again, a is the number of recruits per spawner at low stock sizes, and b is a parameter controlling the strength of density dependence - when $b = 0$, this model *also* collapses to the density-independent model. The peak recruitment, which is given by $R_p = a/be$, occurs at an intermediate stock size ($1/b$).

Let's plot the three of these models together to get an idea for what they look like. Play around with the values of a and b - what do you notice about the shape of the Beverton-Holt and Ricker stock-recruitment curves as you vary the parameters?

> **As we increase the value of a, the slope and maximum number of recruits increases for all of these functions. As we increase the value of b, the Beverton-Holt model reaches an asymptote more quickly, and the Ricker model reaches a recruitment peak at lower stock sizes, past which recruitment drops off.**

```r
a <- 0.5
b <- 0.02

sr_curves <- tibble(
  stock = seq(0, max(LakeTroutGIS$stock), length.out = 100),
  linear = a * stock,
  `Beverton-Holt` = (a * stock)/(1 + b * stock),
  Ricker = (a * stock) * (exp(-b * stock))
)

sr_curves %>%
  pivot_longer(cols = -stock, names_to = "form", values_to = "recruits") %>%
  ggplot(aes(stock, recruits))  +
  geom_line() +
  facet_wrap(~form, scales = "free_y")
```



Likelihood functions

It’s typical of recruitment data for the variance to increase with the mean, because errors are often *multiplicative* instead of additive. Why might this be?

> **If we think of the sources of variability in this data, such as environmental stochasticity, as altering the per-capita survival of juvenile fish or the fecundity of adult fish in the population, than “good” and “bad” years multiply the number of surviving offspring, instead of simply resulting in more or less offspring surviving.**

Thus far, when we’ve used a normal likelihood function, we’ve assumed that the variance is constant (i.e., does not depend on the value of the predictors) - in this case, not only do we expect the variance to be greater for larger recruitment values, but we also know that recruitment can’t be negative, so if our fitted recruitment value for stock size $S = 0$ is also 0 (i.e., $E(R|S = 0) = 0$), than our recruitment at $S = 0$ would be $R|S = 0 \sim \text{Normal}(0, \sigma^2)$, and half of the probability mass would be on recruitment values that aren’t possible.

We can solve this a number of ways - we could choose a probability distribution designed for non-negative, skewed data, such as the lognormal or Gamma distributions. Or, we could just log-transform our recruitment values and the corresponding stock-recruitment functions, and use a Normal distribution. For each of these models, this gives:

$$
\begin{aligned}
\log(R) &\sim \text{Normal}(\log(\mu), \sigma^2) \\
\log(\mu) &= \log(aS) && \text{(linear)} \\
\log(\mu) &= \log\left(\frac{aS}{1+bS}\right) && \text{(Beverton-Holt)} \\
\log(\mu) &= \log\left(aSe^{-bS}\right) && \text{(Ricker)}
\end{aligned}
$$

I've written the first of these for you - you write the other two. The only constraints on these models are that $a, b >= 0$.

```
## Negative log-likelihood for density-independent model
linear_nll <- function(par, s, r) {
  a <- par[1]; sigma <- par[2]
  if (a >= 0) {
    mu_r <- log(a * s)
    nll <- -sum(dnorm(log(r), mean = mu_r, sd = sigma, log = TRUE))
    return(nll)
  } else {
    return(5000)
  }
}

## Negative log-likelihood for Beverton-Holt model
## mostly the same, but we've added another parameter
## and we've changed the equation for mu_r
beverton_nll <- function(par, s, r) {
  a <- par[1]; b <- par[2]; sigma <- par[3]
  if (a >= 0 & b >= 0) {
    mu_r <- log((a * s)/(1 + b * s))
    nll <- -sum(dnorm(log(r), mean = mu_r, sd = sigma, log = TRUE))
    return(nll)
  } else {
    return(5000)
  }
}

## Negative log-likelihood for Ricker model
## same as the beverton-holt nll function,
## but with a different equation for mu_r
ricker_nll <- function(par, s, r) {
  a <- par[1]; b <- par[2]; sigma <- par[3]
  if (a >= 0 & b >= 0) {
    mu_r <- log((a * s) * (exp(-b * s)))
    nll <- -sum(dnorm(log(r), mean = mu_r, sd = sigma, log = TRUE))
    return(nll)
  } else {
    return(5000)
  }
}
```

Initial values

recruit-to-spawner ratio - a

We can obtain initial values for a the first model using a linear regression (in fact, this would give our MLE for a if we assumed a Normal distribution). We'll store this value, and the initial values for a in the other two models, to a list. We can fix the intercept at 0 by including a 0 in our formula:

```
sr_lm <- lm(recruits ~ 0 + stock, data = LakeTroutGIS)
a_init <- list("linear" = as.numeric(coef(sr_lm)))
```

For the a coefficient in the Beverton-Holt and *Ricker* models, one way that we can obtain reasonable starting values by fitting the same regression to the data at small stock sizes. This definitely isn't the best way - the `FSA` package provides a less shoddy method with the `srStarts` function, but it's intuitive and it's easy:

```
sr_lm0 <- lm(recruits ~ 0 + stock, data = LakeTroutGIS[LakeTroutGIS$stock < 30, ])
a_init$beverton <- a_init$ricker <- as.numeric(coef(sr_lm0))
```

Printing our three (really two) starting values:

```
unlist(a_init)
```

```
##    linear    ricker  beverton
## 0.2787063 0.3928351 0.3928351
```

density-dependent parameter - b

For the Beverton-Holt model, we know that the asymptote occurs at a/b. To my eyes, the asymptote occurs at about 20, so this gives a starting value for b of $a/20 \approx 0.06$.

For the Ricker model, the peak recruitment (which to me looks to be about 25) occurs at $R_p = a/be$ (so $b = a/(R_p e)$), so an initial value for b could be $b = a/(25 \times e) \approx 0.006$.

Let's store these in a list:

```
b_init <- list(
  beverton = a_init$beverton/20,
  ricker = a_init$ricker/(25 * exp(1))
)

unlist(b_init)
```

```
##    beverton      ricker
## 0.019641755 0.005780638
```

Sigma

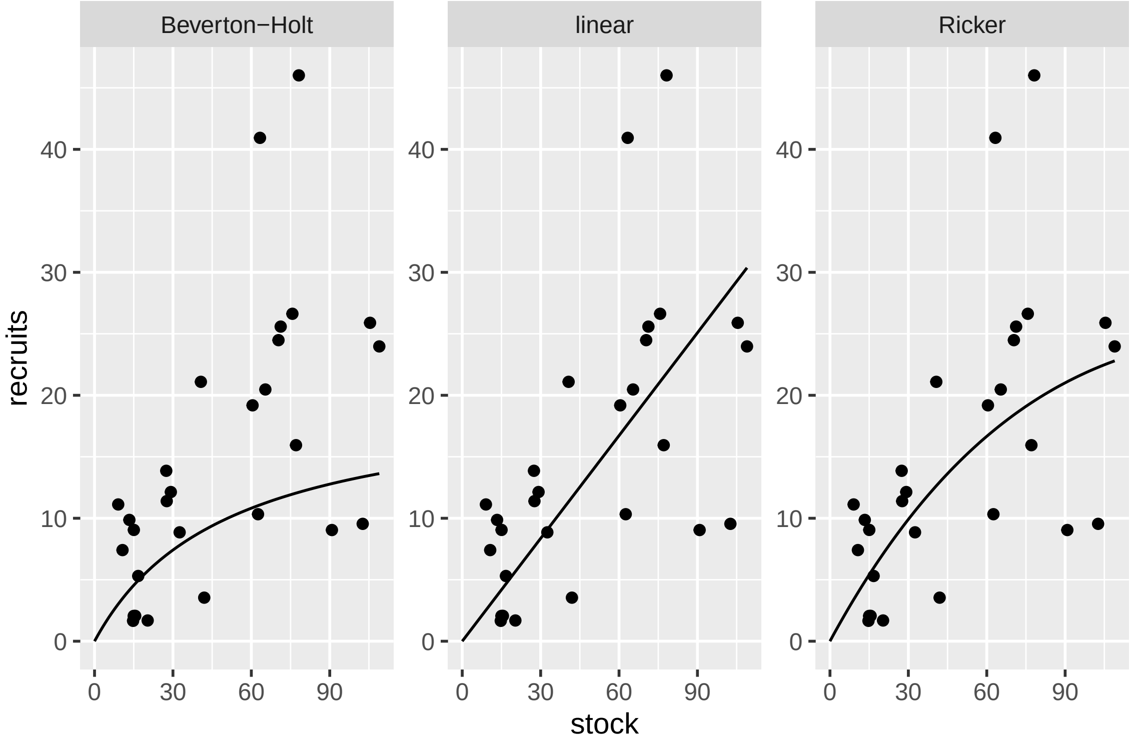
Finally, we need an initial value for the standard deviation of the residuals. Of course, a good starting value is... the standard deviation of the residuals. So, taking the log of the recruitment and stock size and fitting a regression:

```
sr_lm_log <- lm(log(recruits) ~ log(stock), data = LakeTroutGIS)
sigma_init <- sd(resid(sr_lm_log))
sigma_init
```

```
## [1] 0.71046
```

Now, let's plot the stock recruitment curves given by our chosen starting values to make sure they're reasonable:

```
sr_curves <- tibble(
  stock = seq(0, max(LakeTroutGIS$stock), length.out = 100),
  linear = a_init$linear * stock,
  `Beverton-Holt` = (a_init$beverton * stock)/(1 + b_init$beverton * stock),
  Ricker = (a_init$ricker * stock) * (exp(-b_init$ricker * stock))
)

sr_curves %>%
  pivot_longer(cols = -stock, names_to = "form", values_to = "recruits") %>%
  ggplot(aes(stock, recruits))  +
  geom_point(data = LakeTroutGIS) +
  geom_line() +
  facet_wrap(~form, scales = "free_y")
```



Maximum-Likelihood Estimates

By now, we know that obtaining the MLE once we've already written down the likelihood functions and obtained starting values is pretty easy:

```
linear_mle <- optim(
  par = c(a = a_init$linear, sigma = sigma_init),
  fn = linear_nll,
  s = LakeTroutGIS$stock,
  r = LakeTroutGIS$recruits,
  hessian = TRUE
)

beverton_mle <- optim(
  par = c(a = a_init$beverton, b = b_init$beverton, sigma = sigma_init),
  fn = beverton_nll,
  s = LakeTroutGIS$stock,
  r = LakeTroutGIS$recruits,
  hessian = TRUE
)

ricker_mle <- optim(
  par = c(a = a_init$ricker, b = b_init$ricker, sigma = sigma_init),
  fn = ricker_nll,
  s = LakeTroutGIS$stock,
  r = LakeTroutGIS$recruits,
  hessian = TRUE
)
```

Printing the fitted values:

```
MLE <- list(
  "linear" = linear_mle$par,
  "Beverton-Holt" = beverton_mle$par,
  "Ricker" = ricker_mle$par
)
MLE
```

```
## $linear
##         a     sigma
## 0.2859496 0.7211662
##
## $`Beverton-Holt`
##           a           b       sigma
## 0.382775616 0.007363847 0.700587266
##
## $Ricker
##           a           b       sigma
## 0.373835046 0.005514563 0.699752160
```

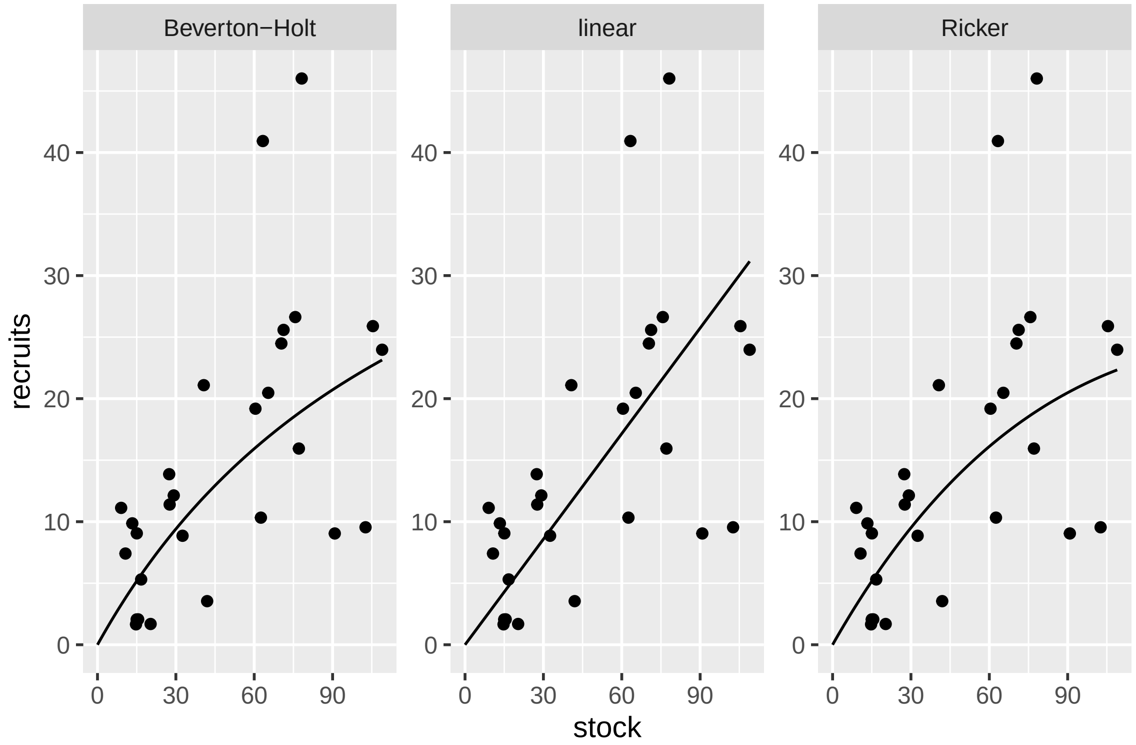
And their standard errors:

```
var_cov <- list(
  "linear" = solve(linear_mle$hessian),
  "Beverton-Holt" = solve(beverton_mle$hessian),
  "Ricker" = solve(ricker_mle$hessian)
)

se <- lapply(var_cov, function(x) sqrt(diag(x)))
se
```

```
## $linear
##          a      sigma
## 0.03897168 0.09636486
##
## $`Beverton-Holt`
##           a           b       sigma
## 0.108646026 0.007620346 0.093622023
##
## $Ricker
##           a           b       sigma
## 0.090675168 0.004180997 0.093506634
```

Let's plot the fitted curves implied by our maximum likelihood estimates for a and b:

```
sr_curves <- tibble(
  stock = seq(0, max(LakeTroutGIS$stock), length.out = 100),
  linear = linear_mle$par[1] * stock,
  `Beverton-Holt` = (beverton_mle$par[1] * stock)/(1 + beverton_mle$par[2] * stock),
  Ricker = (ricker_mle$par[1] * stock) * (exp(-ricker_mle$par[2] * stock))
)

sr_curves %>%
  pivot_longer(cols = -stock, names_to = "form", values_to = "recruits") %>%
  ggplot(aes(stock, recruits))  +
  geom_point(data = LakeTroutGIS) +
  geom_line() +
  facet_wrap(~form, scales = "free_y")
```



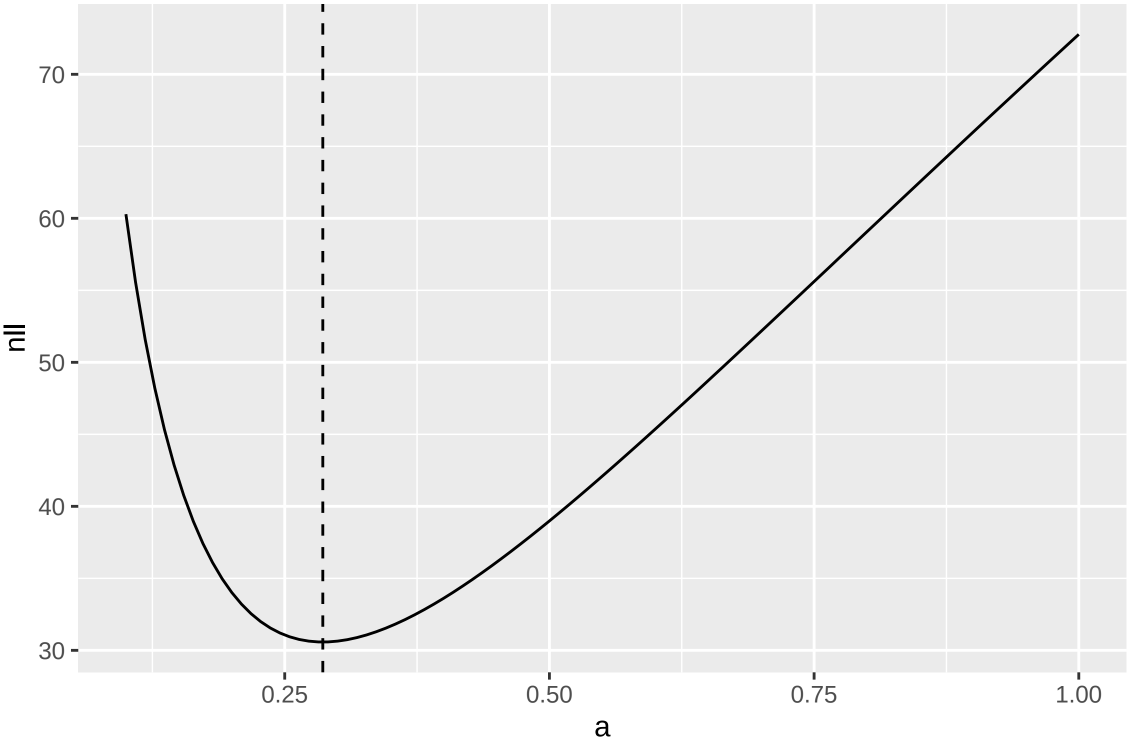
Even though the underlying form of the density-dependence differs between the Beverton-Holt and Ricker models, the best-fit models produce very similar curves.

Likelihood Slices

Here are likelihood slices for each of these models - for each model, I've plotted the negative log likelihood as a function of the parameter values, while holding σ at its maximum likelihood estimate.

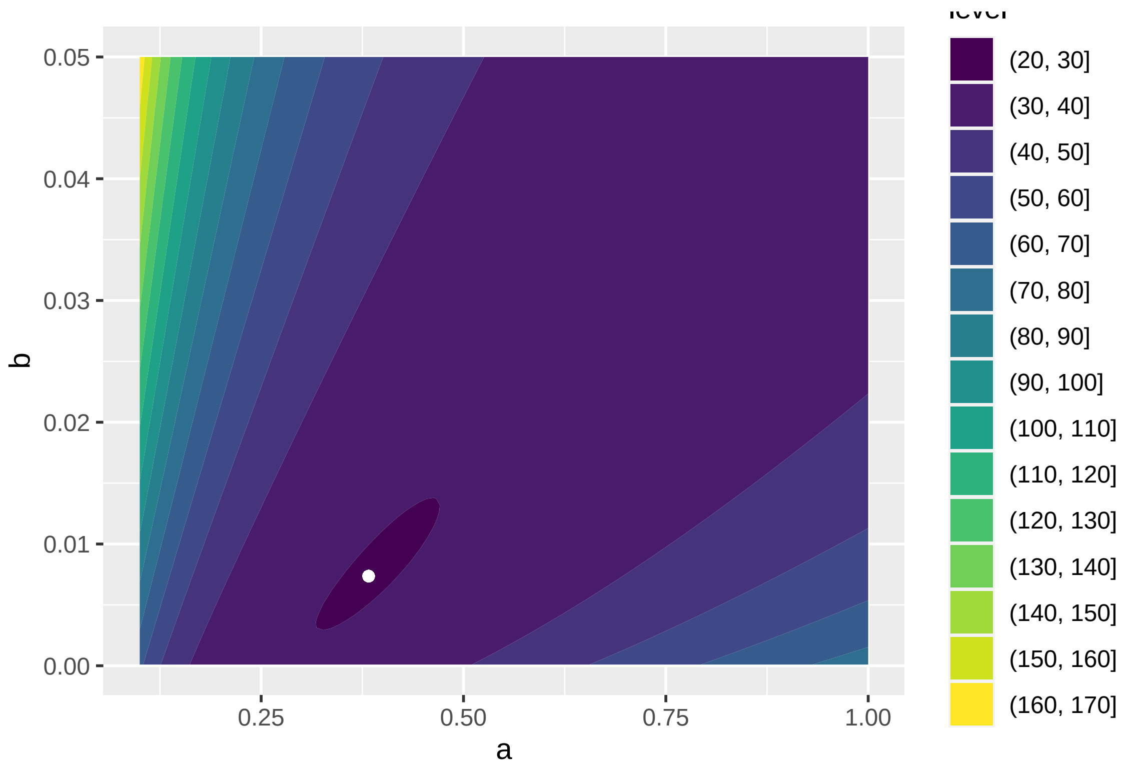
Density-Independent Model

```
a_seq <- seq(0.1, 1, length.out = 100)
nll_linear <- rep(NA, length(a_seq))

for (i in 1:length(a_seq)) {
  nll_linear[i] <- linear_nll(
    c(a_seq[i], sigma = linear_mle$par[2]),
    s = LakeTroutGIS$stock,
    r = LakeTroutGIS$recruits
  )
}

data.frame(a = a_seq, nll = nll_linear) %>%
  ggplot(aes(a, nll)) +
  geom_line() +
  geom_vline(aes(xintercept = linear_mle$par[1]), linetype = "dashed")
```



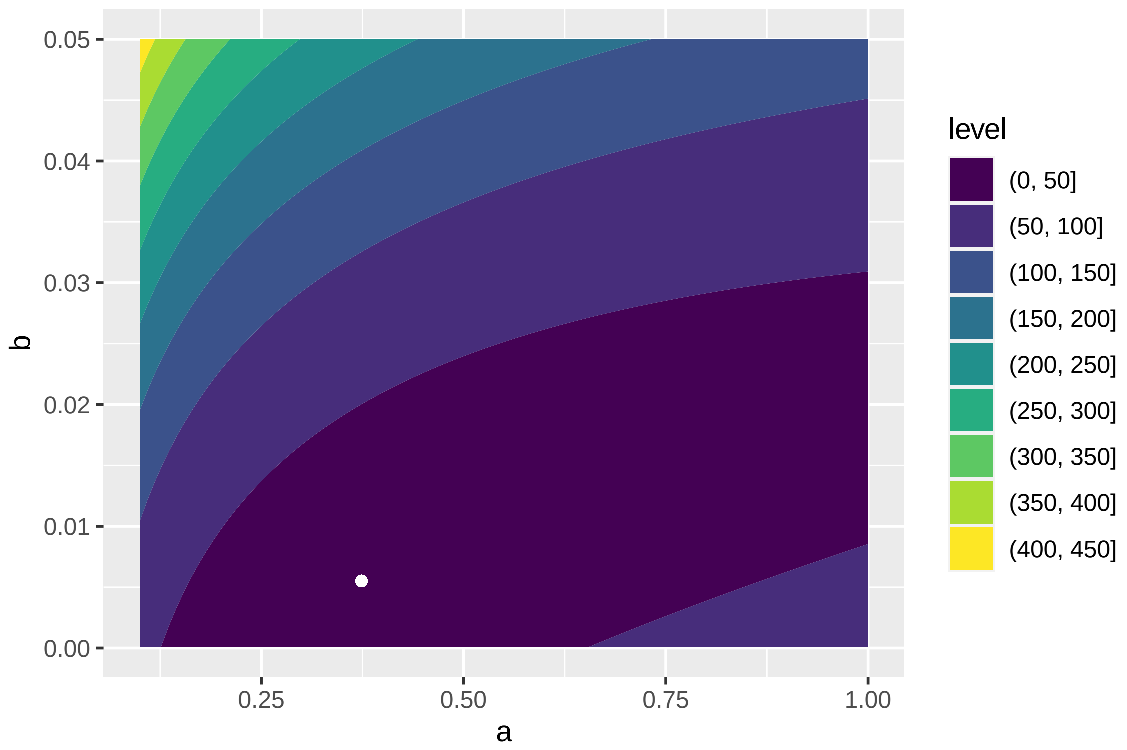
Beverton-Holt Model

```
a_seq <- seq(0.1, 1, length.out = 100)
b_seq <- seq(0.0001, 0.05, length.out = 100)
par_grid <- expand.grid(a = a_seq, b = b_seq)

for (i in 1:nrow(par_grid)) {
  par_grid$nll_beverton[i] <- beverton_nll(
    c(par_grid$a[i], par_grid$b[i], sigma = beverton_mle$par[3]),
    s = LakeTroutGIS$stock,
    r = LakeTroutGIS$recruits
  )
}

par_grid %>%
  ggplot(aes(a, b, z = nll_beverton)) +
  geom_contour_filled() +
  geom_point(aes(x = beverton_mle$par[1], y = beverton_mle$par[2]),
             color = "white")
```



Ricker Model

```
for (i in 1:nrow(par_grid)) {
  par_grid$nll_ricker[i] <- ricker_nll(
    c(par_grid$a[i], par_grid$b[i], sigma = ricker_mle$par[3]),
    s = LakeTroutGIS$stock,
    r = LakeTroutGIS$recruits
  )
}

par_grid %>%
  ggplot(aes(a, b, z = nll_ricker)) +
  geom_contour_filled() +
  geom_point(aes(x = ricker_mle$par[1], y = ricker_mle$par[2]),
             color = "white")
```



Exercises

1. Confidence intervals

Do you think asymptotic (i.e. Normal) confidence intervals would be a good choice for the a and b parameters in the Beverton-Holt and Ricker models? In answering this question, think about (1) the maximum-likelihood estimates, (2) the Fisher standard errors on those estimates, and (3) the shape of the likelihood slices we plotted. Note: I'm not actually asking you to compute confidence intervals for this question.

> **If the distribution of the maximum likelihood estimates for a and b were approximately normal, we would expect the likelihood slices to be relatively symmetrical around the maximum likelihood estimates, with contours that form ellipses as they did in the first maximum likelihood lab on linear regression. Additionally, the standard errors for the b terms in both the Beverton-Holt and Ricker models are nearly as large as the estimates themselves, so if we were to construct normal confidence intervals for b, they would include values of b below zero, which doesn't seem to make sense. So, we'd probably want to use bootstrap or likelihood profile confidence intervals here.**

2. Model Selection

Compute AIC for each of these models. To make this quicker, I've defined a function below that takes the output of `optim` and computes and AIC value. Based on AIC, is there any evidence for density-dependence in the stock-recruitment relationship?

> **The AIC values are all pretty similar for these models, which means that the Beverton-Holt and Ricker models don't fit the data all that much better than the density-independent model. We could interpret this as a lack of evidence for density-dependence in the stock-recruitment relationship**

```
AIC <- function(opt) 2*length(opt$par) + 2*opt$value

model_AIC <- list(
  linear = AIC(linear_mle),
  beverton = AIC(beverton_mle),
  ricker = AIC(ricker_mle)
)

model_AIC
```

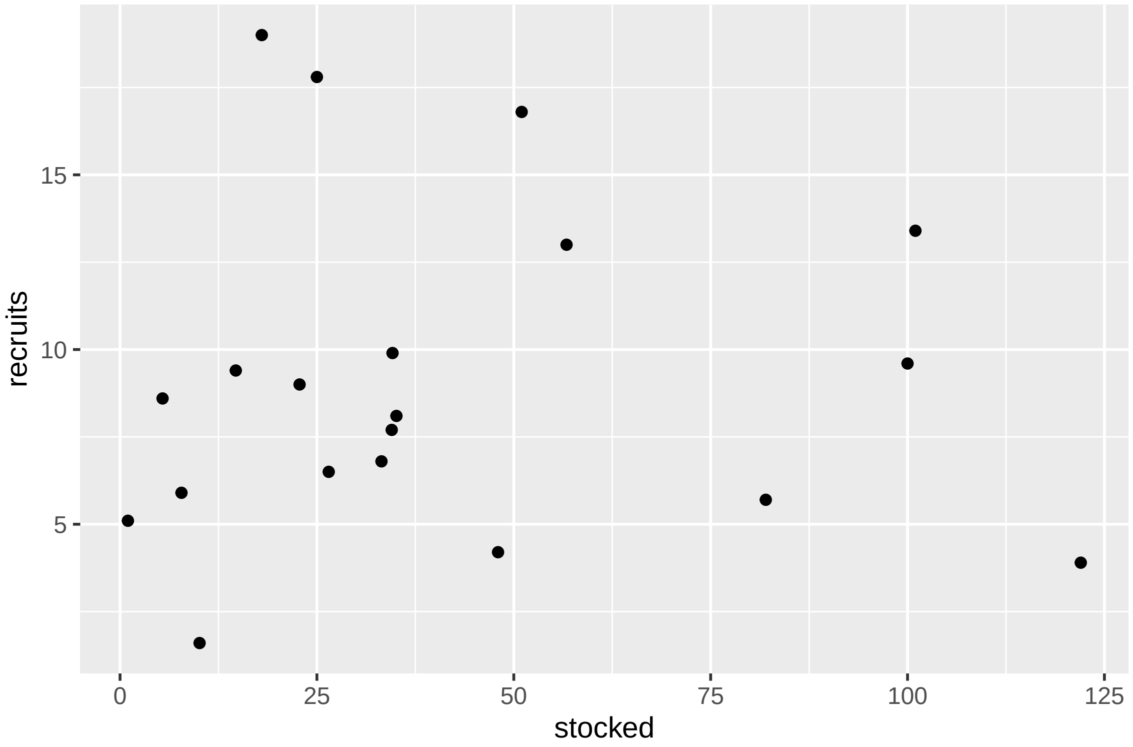
```
## $linear
## [1] 65.15658
##
## $beverton
## [1] 65.53279
##
## $ricker
## [1] 65.46726
```

3. Likelihood-ratio tests

We can also assess whether there's density-dependence or not using a likelihood ratio test, by fixing the parameter b in either the Beverton-Holt or Ricker model to zero (in which case, the model collapses to the density-independent model), and then computing a likelihood-ratio test.

Let's load a dataset containing stock recruitment data for a different part of lake Superior, and conduct this test with the Ricker model, using the `linear_nll` function to fit our density-independent model.

```
MI7 <- LakeTroutMI %>% filter(area == "MI7" & !is.na(recruits))
MI7 %>% ggplot(aes(stocked, recruits)) + geom_point()
```



We'll just shortcut the whole process of choosing parameter values for a and b by using the `srStarts` function from the `FSA` ("fisheries stock assessment") package:

```
## initial values for a and b
ricker_init <- unlist(srStarts(recruits ~ stocked, data = MI7, type = "Ricker"))
```

and again we'll choose an initial value for sigma with linear regression:

```
MI7_lm <- lm(log(recruits) ~ 0 + log(stocked), data = MI7)

## vector of initial parameters for linear fit
linear_init <- c(a = as.numeric(coef(MI7_lm)), sigma = sd(resid(MI7_lm)))

## add sigma to initial values for ricker model
ricker_init[["sigma"]] <- sd(resid(MI7_lm))
```

Now, fit the Ricker and density-independent models and compute the likelihood ratio test using the procedure given in the exercises from the first lab. Is there evidence for density-dependence in the stock-recruitment relationship for this population at the $\alpha = 0.05$ level?

> **We can obtain maximum-likelihood estimates using the `ricker_nll` and `linear_nll` functions we've already defined:**

```
ricker_mle <- optim(ricker_init, ricker_nll, s = MI7$stocked, r = MI7$recruits)
linear_mle <- optim(linear_init, linear_nll, s = MI7$stocked, r = MI7$recruits)
```

> **Next, we can obtain the χ^2 test statistic by taking twice the difference of then negative log-likelihoods:**

```
chisq <- 2*(linear_mle$value - ricker_mle$value)
```

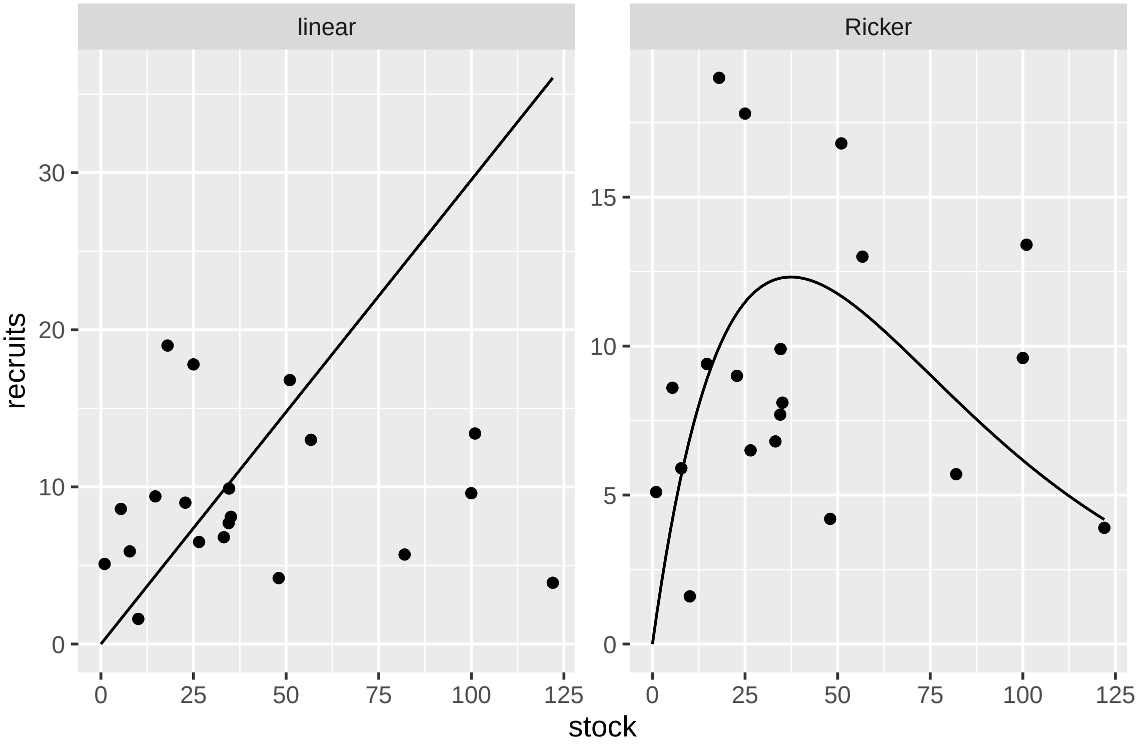
> **We can obtain a p-value for the likelihood-ratio test by comparing this value to the χ_2 distribution with degrees of freedom equal to the difference in the number of parameters between these models. The density-independent model has 2 parameters (a and σ), and the Ricker model has three (a, b, and σ), so the degrees of freedom for our test is 1:**

```
pchisq(chisq, df = 1, lower.tail = FALSE)
```

```
## [1] 9.174419e-06
```

> **Since the p-value is less than our chosen α value of 0.05, the Ricker model fits this data significantly better than does the density-independent model, which we can consider significant evidence for density-dependence in the stock recruitment relationship for this data. Let's plot the fitted lines from these models to see how they compare:**

```
sr_curves <- tibble(
  stock = seq(0, max(MI7$stocked), length.out = 100),
  linear = linear_mle$par[1] * stock,
  Ricker = (ricker_mle$par[1] * stock) * (exp(-ricker_mle$par[2] * stock))
)

sr_curves %>%
  pivot_longer(cols = -stock, names_to = "form", values_to = "recruits") %>%
  ggplot(aes(stock, recruits))  +
  geom_point(aes(x = stocked, y = recruits), data = MI7) +
  geom_line() +
  facet_wrap(~form, scales = "free_y")
```


linear
Ricker
recruits
stock
0
10
20
30
0
5
10
15
0
25
50
75
100
125
0
25
50
75
100
125