ciTools: Quantifying Uncertainty in R

Matthew Avery

Institute for Defense Analyses

John Haman

Bowling Green State University

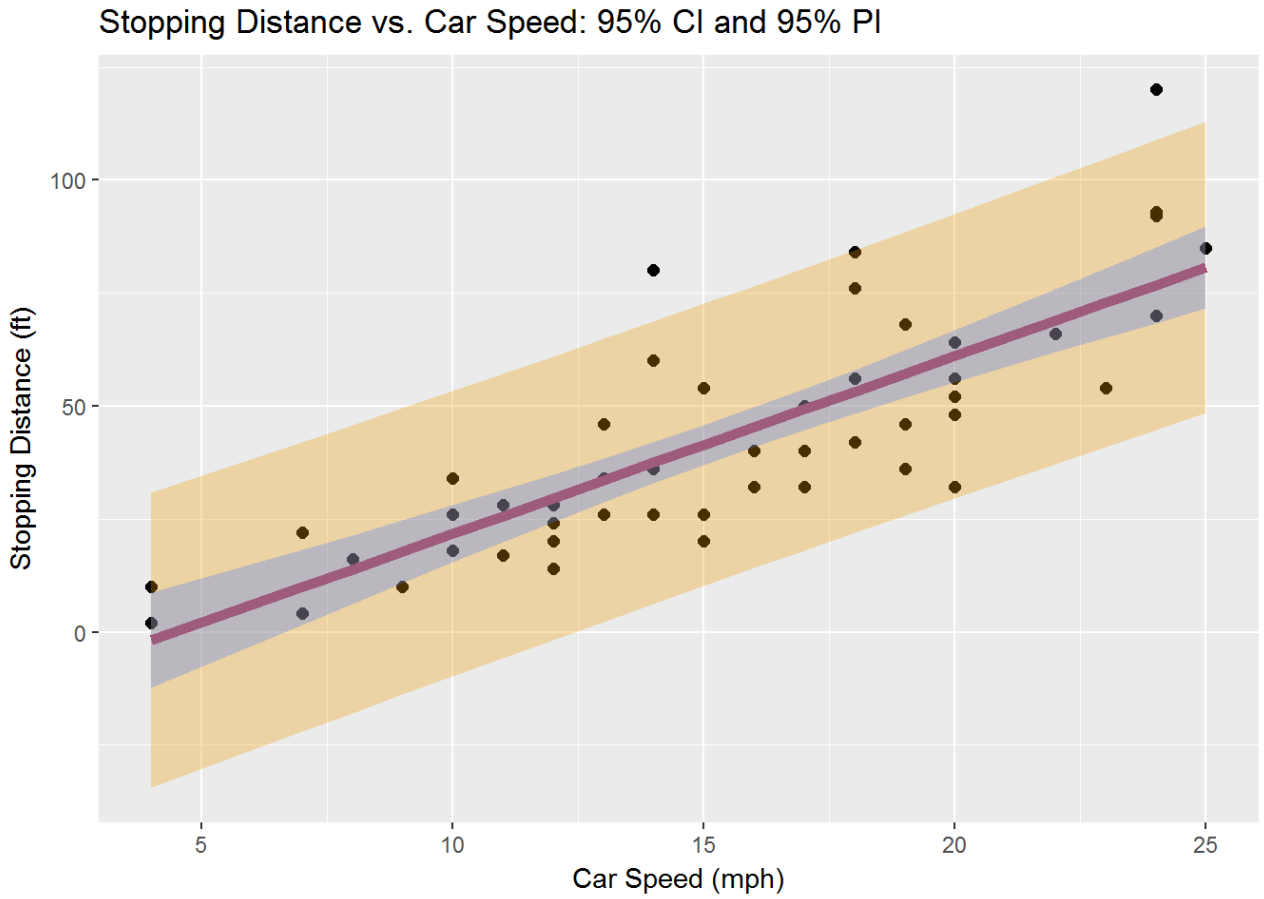
1 February 2018

Use ciTools to add uncertainty estimates to your tibble

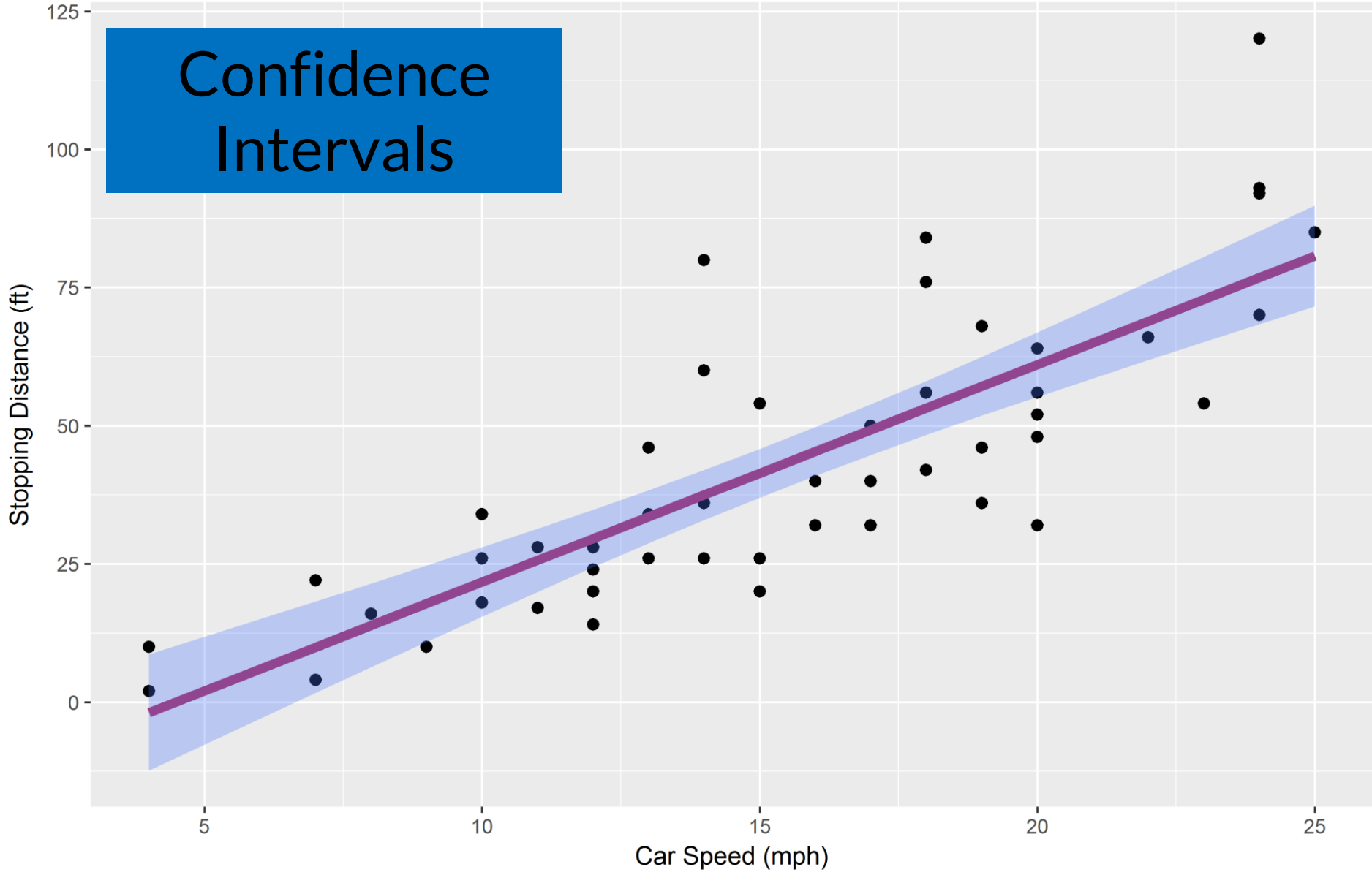
```
library(ciTools)
my_data %>%
  add_ci(model, names = c("lcb", "ucb"))
```

```
## # A tibble: 50 x 5
##    speed  dist      pred        lcb       ucb
##    <dbl> <dbl>     <dbl>      <dbl>     <dbl>
##  1     4     2 -1.849460 -12.329543  8.630624
##  2     4    10 -1.849460 -12.329543  8.630624
##  3     7     4  9.947766   1.678977 18.216556
##  4     7    22  9.947766   1.678977 18.216556
##  5     8    16 13.880175   6.307527 21.452823
##  6     9    10 17.812584  10.905120 24.720047
##  7    10    18 21.744993  15.461917 28.028068
##  8    10    26 21.744993  15.461917 28.028068
##  9    10    34 21.744993  15.461917 28.028068
## 10    11    17 25.677401  19.964525 31.390278
## # ... with 40 more rows
```

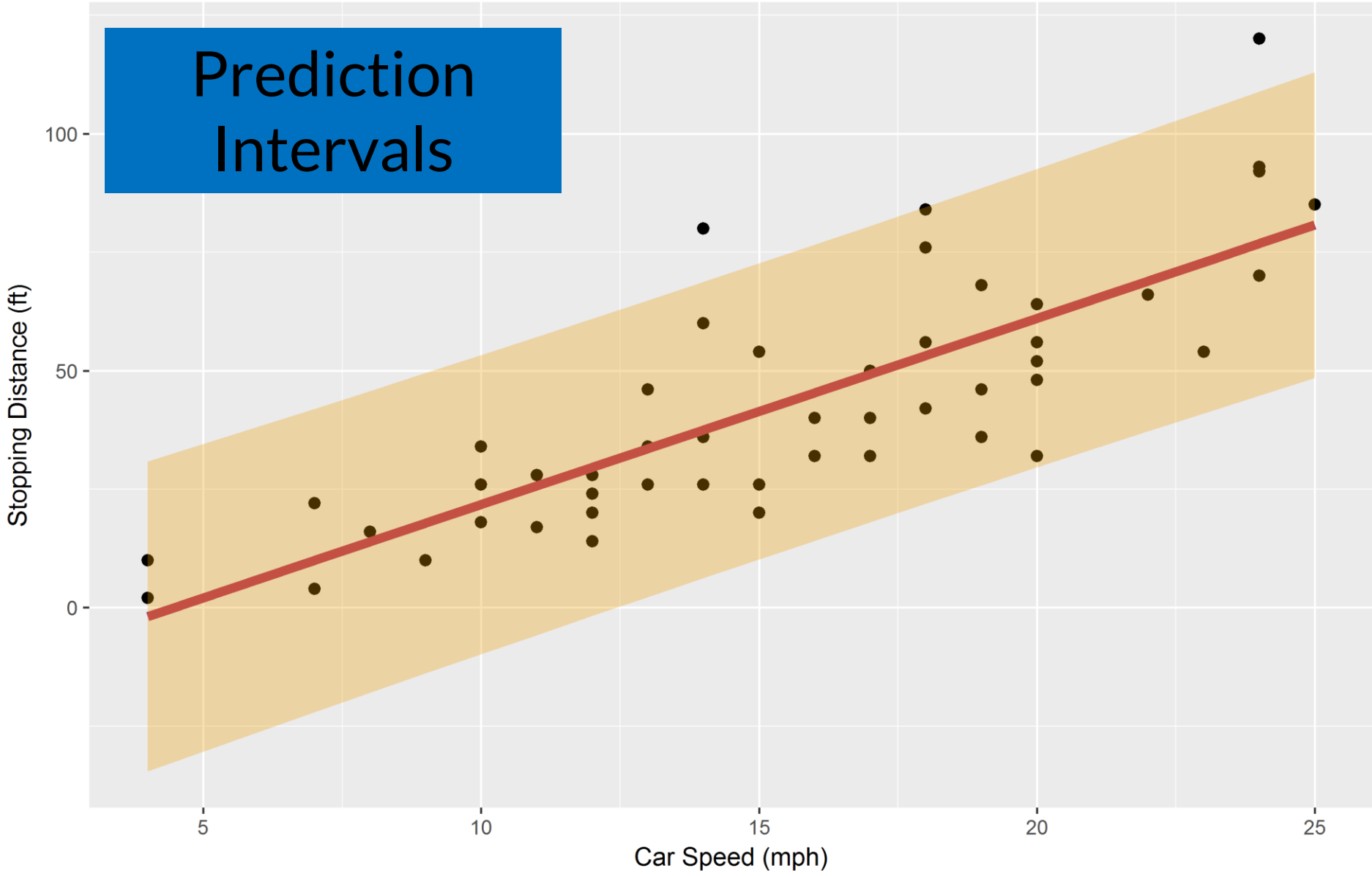
Quickly generate uncertainty intervals, quantiles estimates, and probabilities based on your fitted model



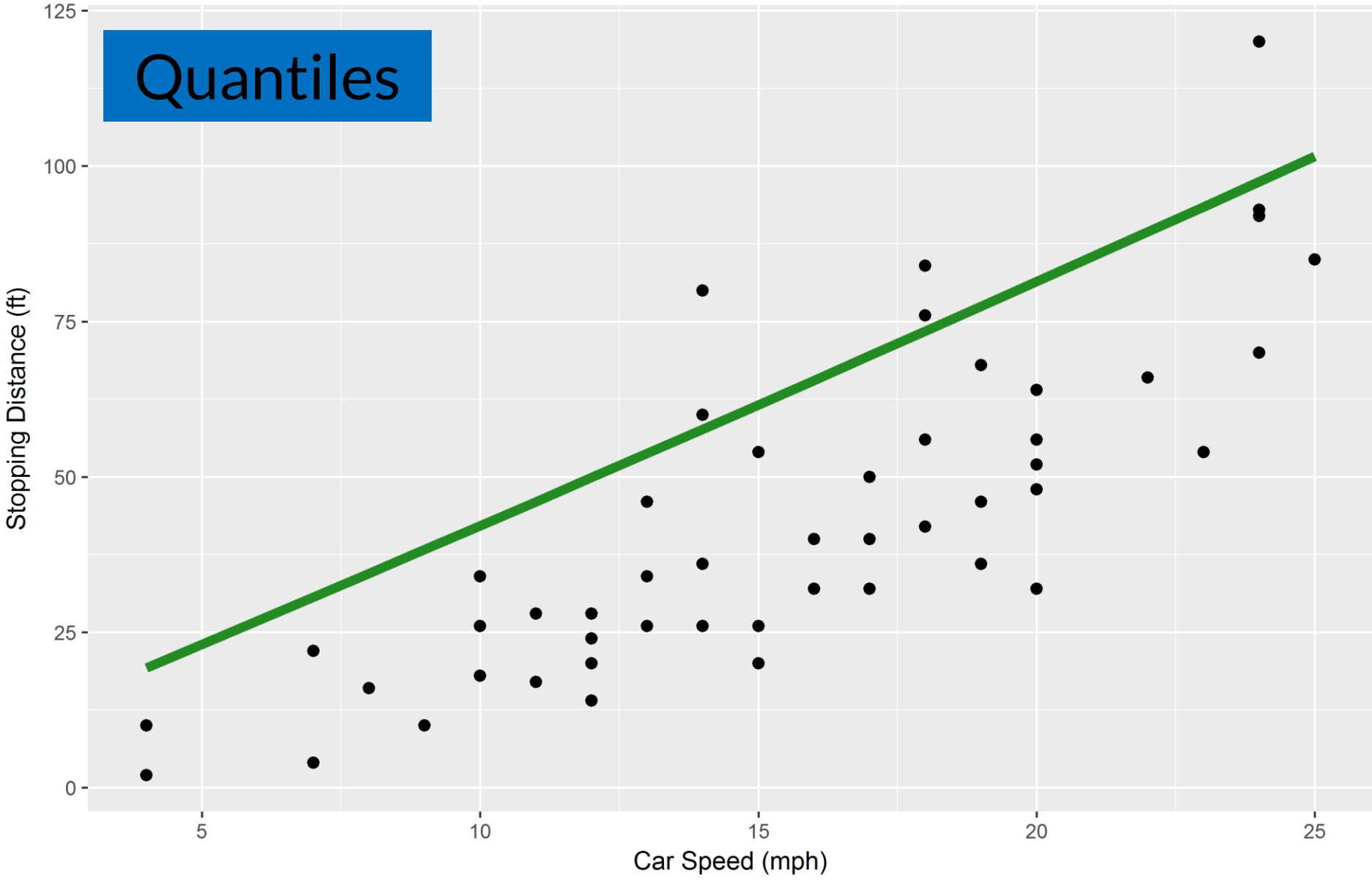
`ciTools::add_ci(tb, fit)`



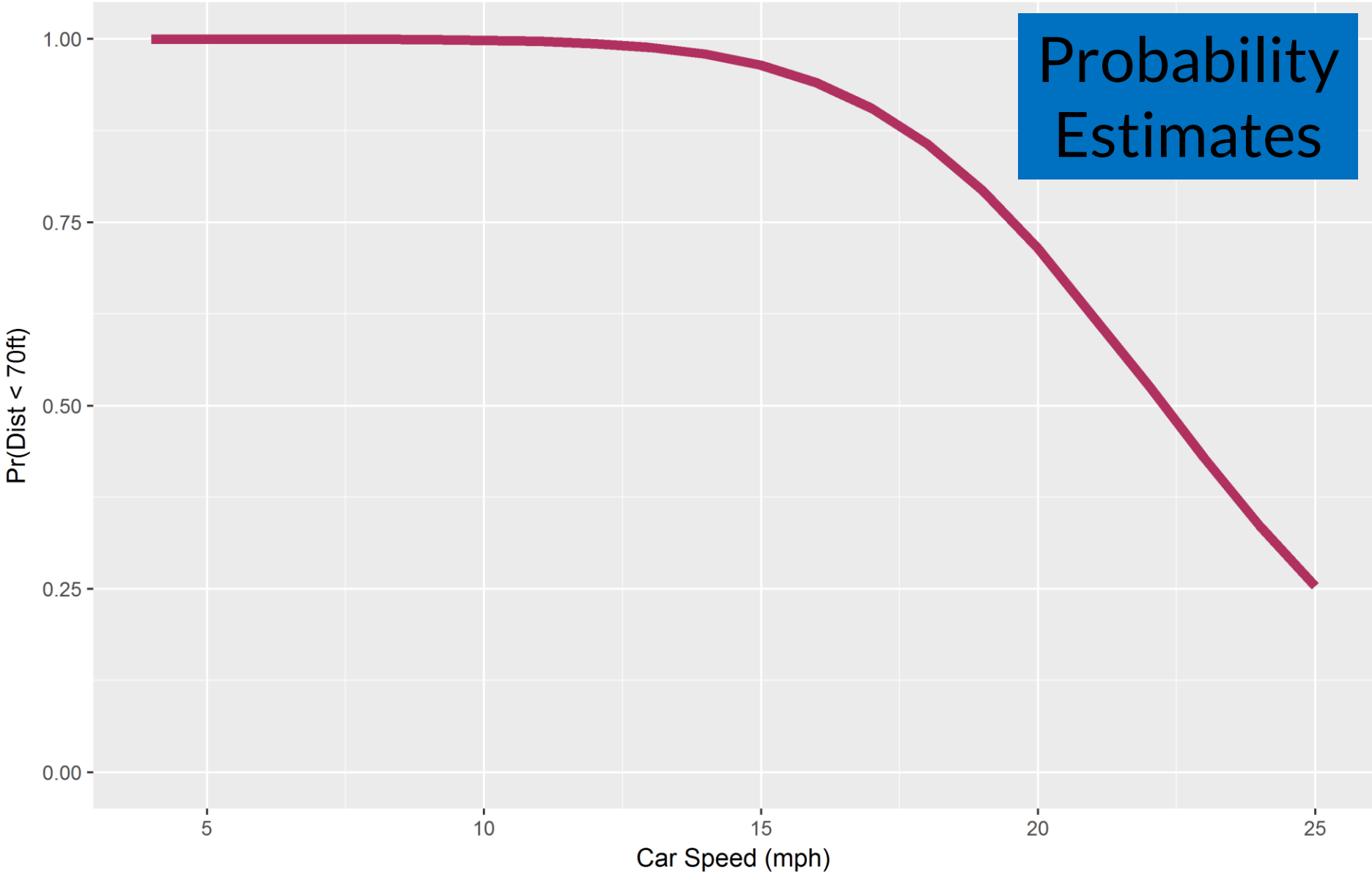
ciTools::add_pi(tb, fit)



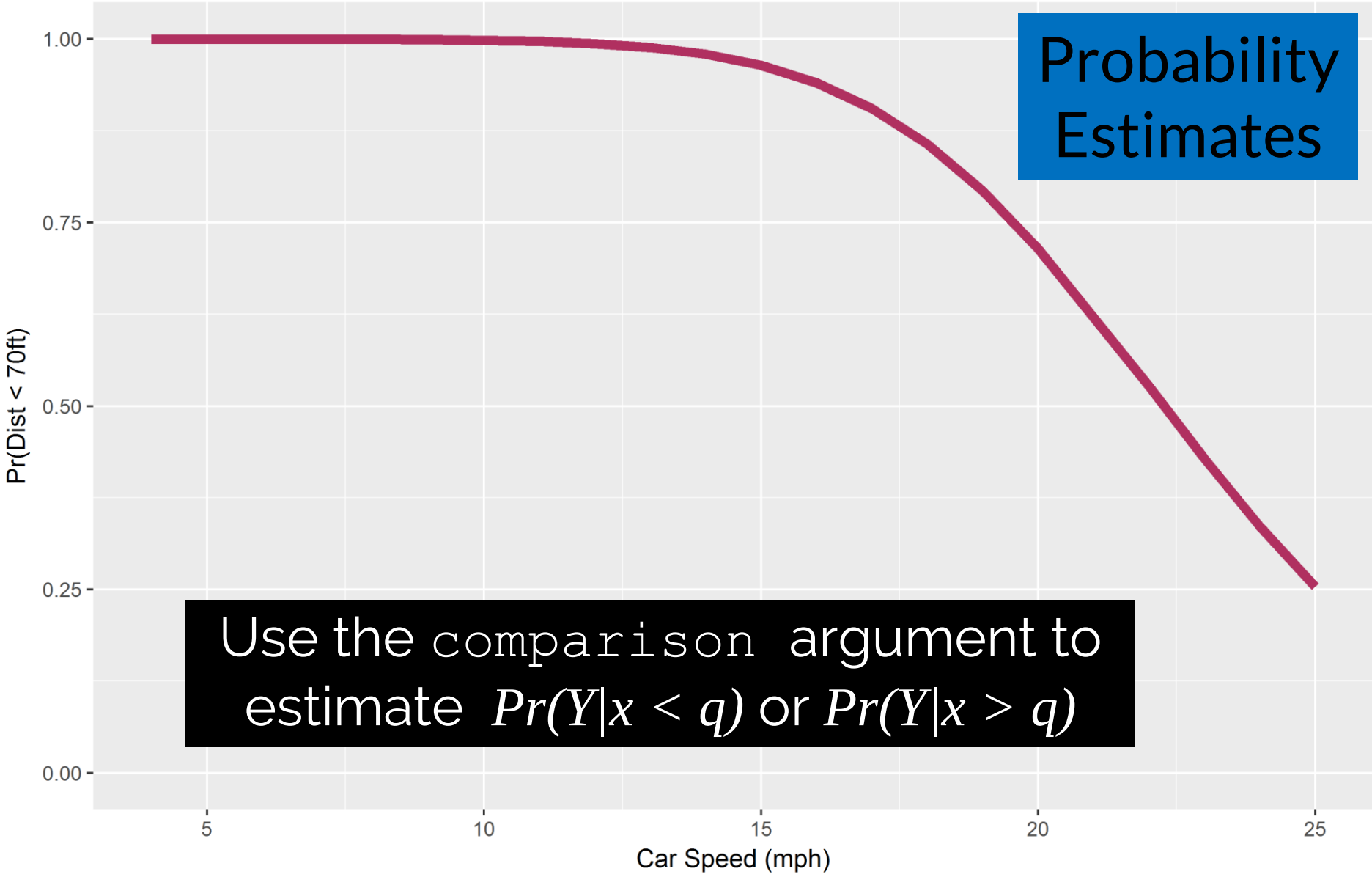
ciTools::add_quantile(tb, fit, q)



ciTools::add_probs(tb, fit, p)



ciTools::add_probs(tb, fit, p)



Uniformity in `ciTools`

`ciTools` works for many types of models, but the syntax doesn't change

Confidence Intervals	**Prediction Intervals**
`add_ci(data, model, …)`	`add_pi(data, model, …)`
Probabilities	**Quantiles**
`add_probs(data, model, quantile, …)`	`add_quantile(data, model, probability, …)`

Inspired by `modelr::add_predictions()`

```
modelr::add_predictions
function (data, model, var = "pred")
{
    data[[var]] <- stats::predict(model, data)
    data
}
```

Powered by generics…

```
add_ci <- function (tb, fit, alpha = 0.05,
      names = NULL, yhatName = "pred", ...) {

UseMethod("add_ci", fit)

}
```

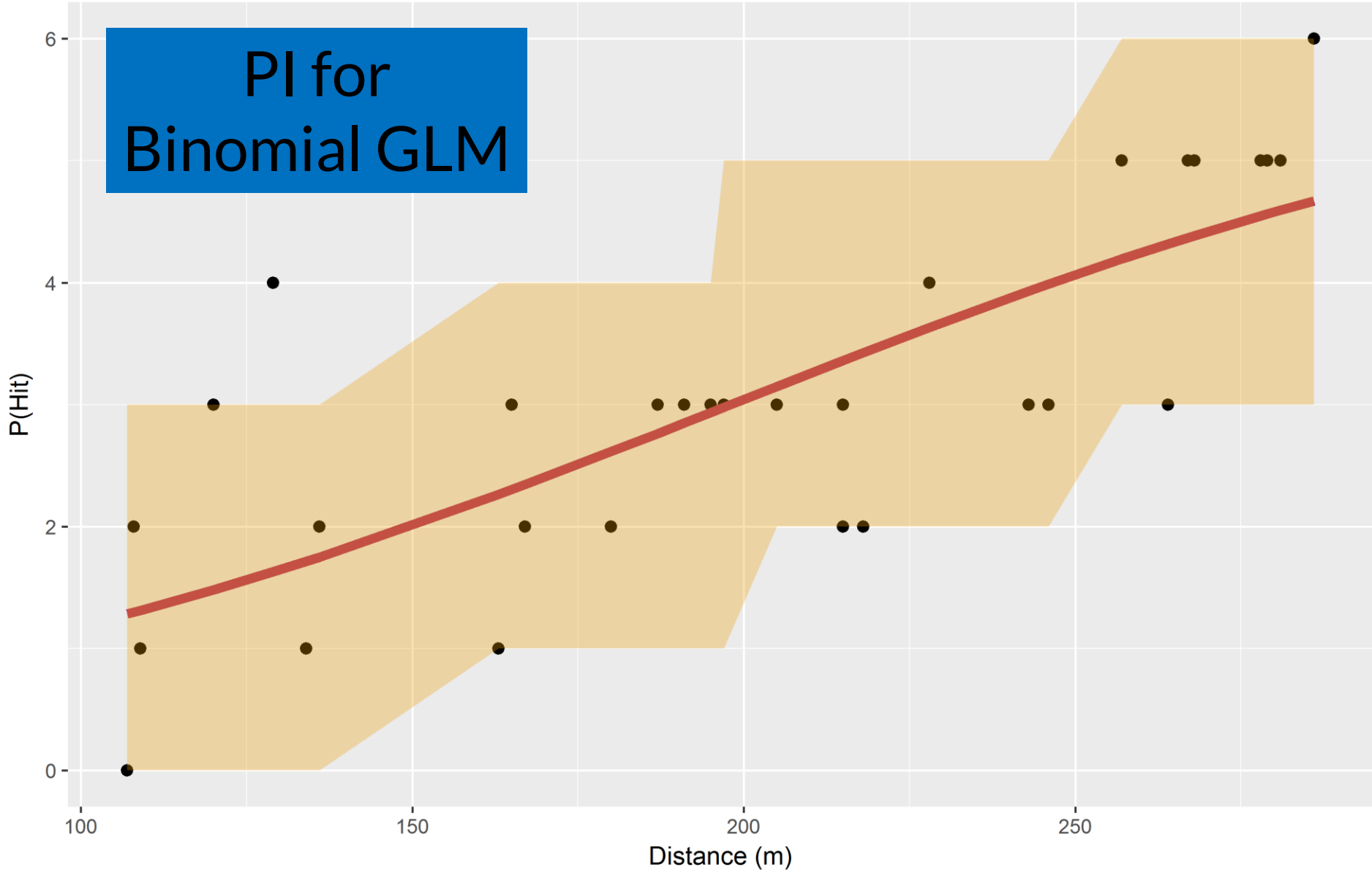
… which choose the right method based on your model

```
add_ci.lm <- function(tb, fit, alpha = 0.05, names = NULL,
      yhatName = "pred", log_response = FALSE, ...){

...

out <- predict(fit, tb, interval = "confidence",
      level = 1 - alpha) tb[[names[1]]] <- out[, 2]
tb[[names[2]]] <- out[, 3]
tb

}
```

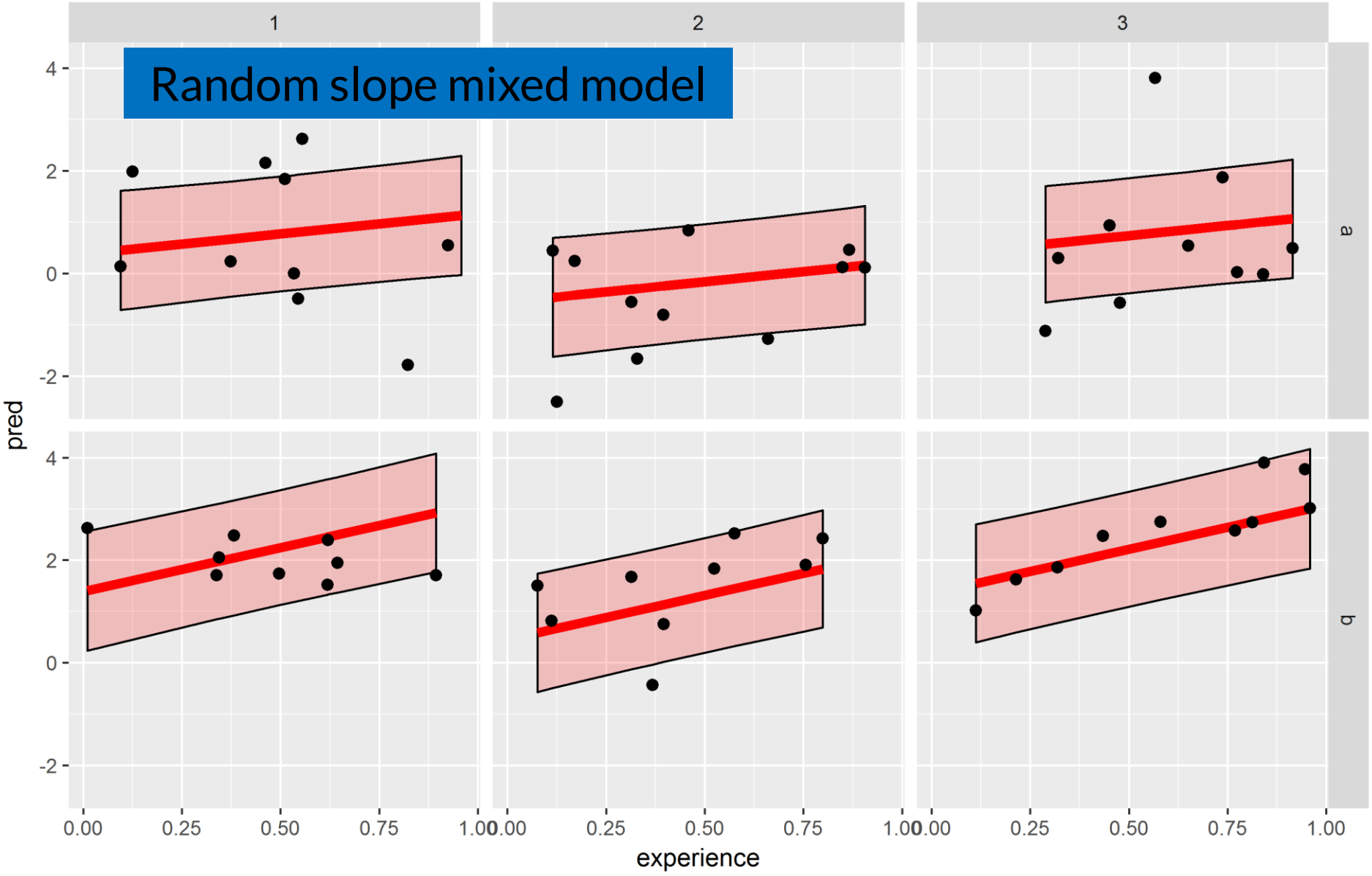
Scope of `ciTools`

R object	Confidence Intervals	Prediction Intervals	Probabilities	Quantiles
`lm`	✓	✓	✓	✓
`lm(ln(y)~...)`	✓	✓	✓	✓
`glm`	✓	✓	✓	✓
`merMod`	✓	✓	✓	✓
`lmer(ln(y)~...)`	In Progress...	✓	✓	✓
`SurvReg`	Future Work	Future Work	Future Work	Future Work
...	...	...	...	...

add_pi.glm(fit, tb, type = “boot”)

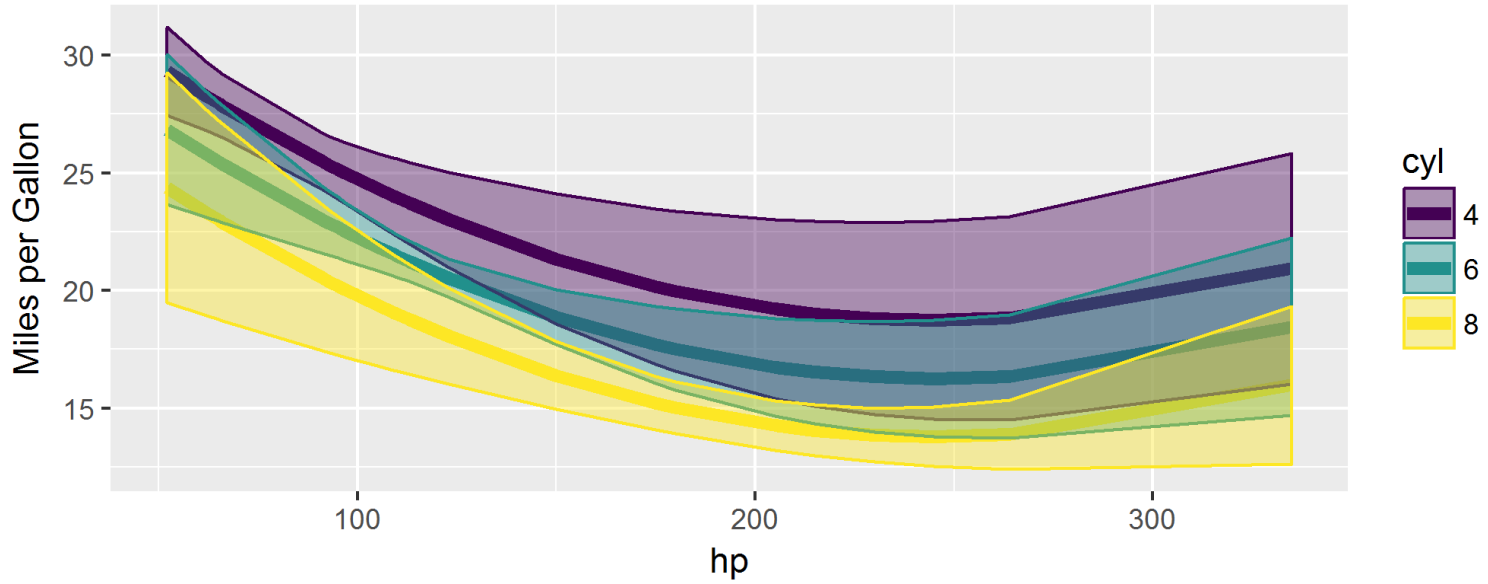


```
add_ci.lmerMod(tb, fit, includeRanef = T)
```



ciTools in the tidy workflow

```
library(tidyverse)
library(ciTools)
library(viridis)

fit <- lm(mpg ~ cyl + hp + I(hp^2) , data = mtcars)
mtcars %>%
  expand(cyl, hp) %>%
  add_ci(fit, alpha = .2, names = c("lower", "upper"), yhatName = "Miles per Gallon") %>%
  mutate(cyl = as.factor(cyl)) %>%
  ggplot(aes(x = hp, y = `Miles per Gallon`, colour = cyl, fill = cyl)) +
  geom_line(size = 2) +
  geom_ribbon(aes(ymin = lower, ymax= upper), alpha = .4) +
  scale_colour_viridis(discrete = T) +
  scale_fill_viridis(discrete = T)
```



Get ciTools where R packages are found!

```
install.packages("ciTools")
```

```
install_github("jthaman/ciTools")
```

Learn how to use ciTools

- Introducing `ciTools`

  CIs, PIs, quantiles, probabilities with lm

- Generalized Linear Models with `ciTools`

  logistic, Poisson, quasi-Poisson methods, simulation studies

- Mixed Models with `ciTools`

  Conditional and unconditional approaches, simulation studies

- Binomial Regression with `ciTools`

  Binomial vs. Bernoulli logistic regression, relevant warning messages