

LAPLACE APPROXIMATION

BAYESIAN STATISTICS FOR ECOLOGISTS

IGB 18. TO 26. NOVEMBER 2019

MCMC METHODS ARE SLOW

- ▶ Markov chains take time to learn the distribution
- ▶ With lots of parameters, memory requirements can be intense
- ▶ Sometimes, fast and approximate is good enough

LAPLACE APPROXIMATION

LAPLACE APPROXIMATION

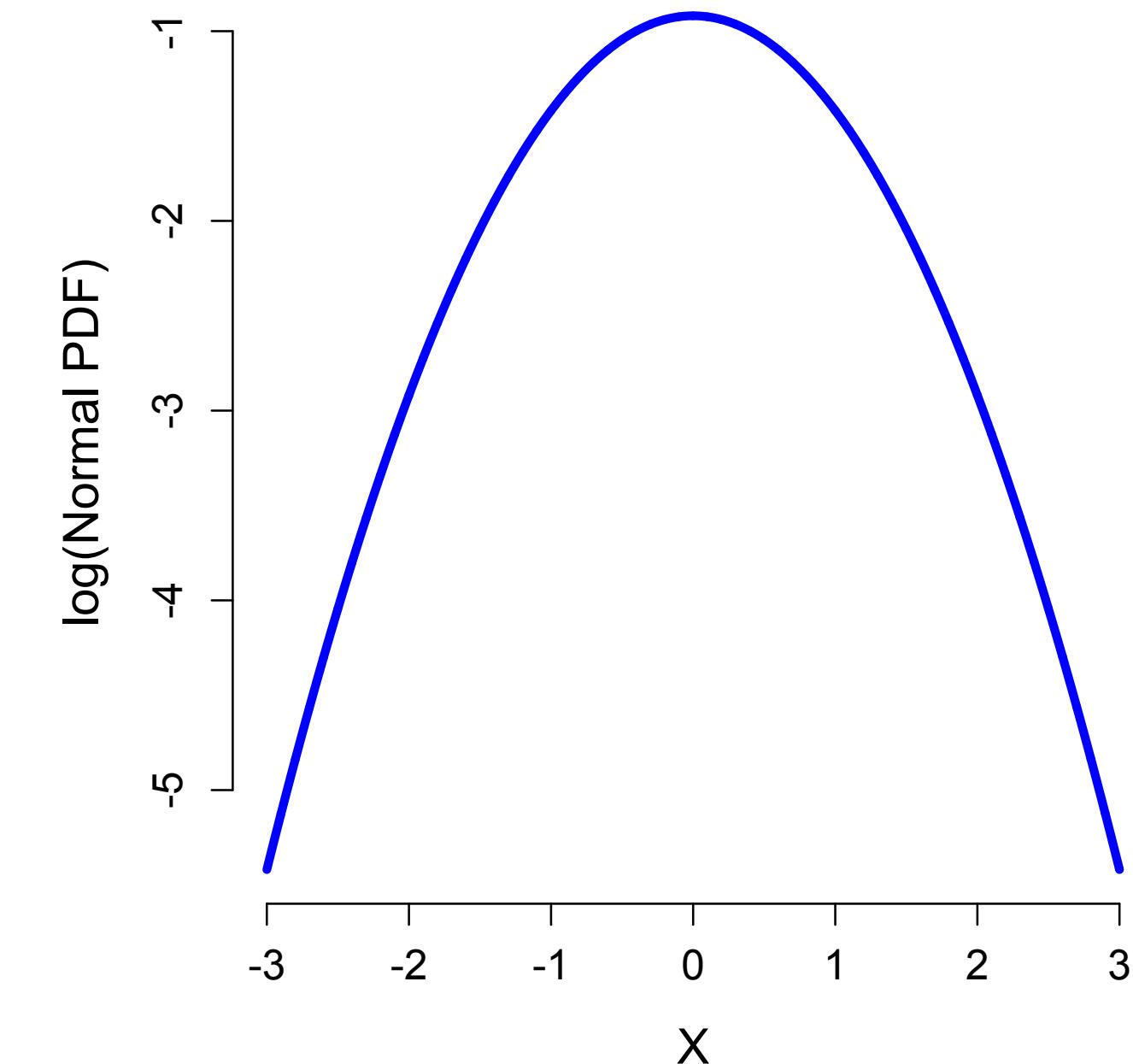
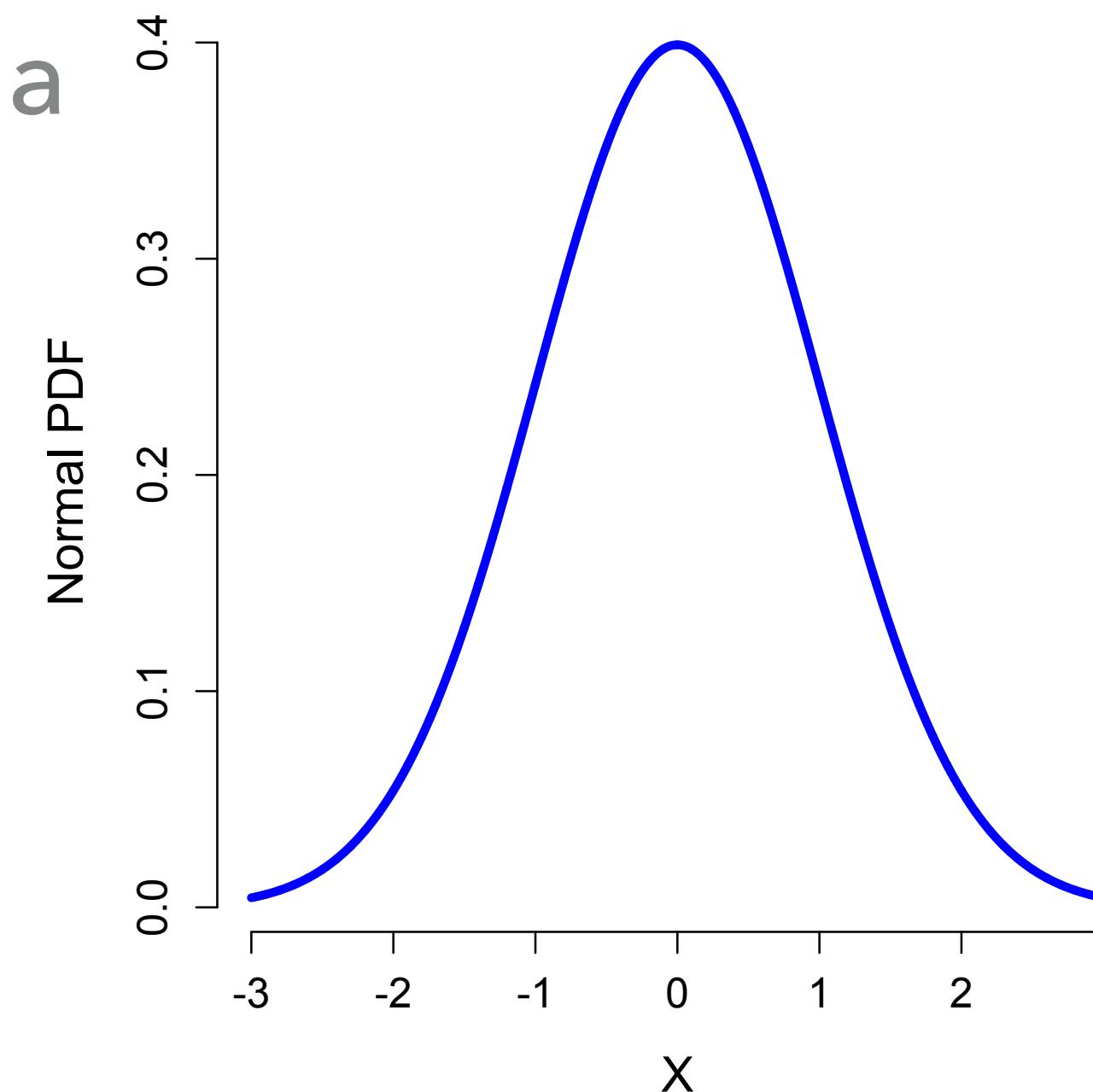
- ▶ This is a general method for approximating the shape of any curve

LAPLACE APPROXIMATION

- ▶ This is a general method for approximating the shape of any curve
- ▶ Because of the central limit theorem, many joint posteriors are approximately multivariate normal distributions

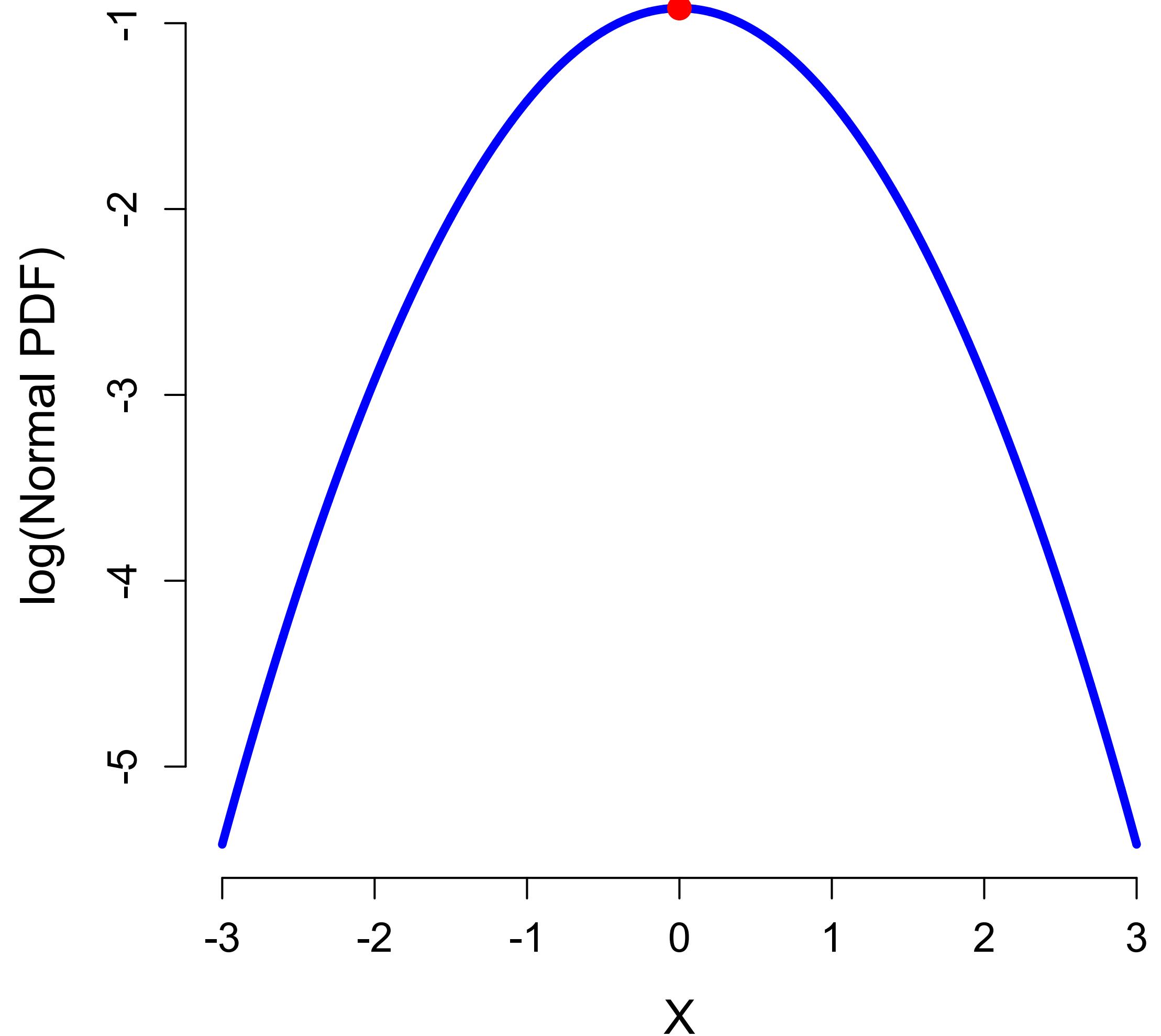
LAPLACE APPROXIMATION

- ▶ This is a general method for approximating the shape of any curve
- ▶ Because of the central limit theorem, many joint posteriors are approximately multivariate normal distributions
- ▶ The log of a gaussian is a parabola



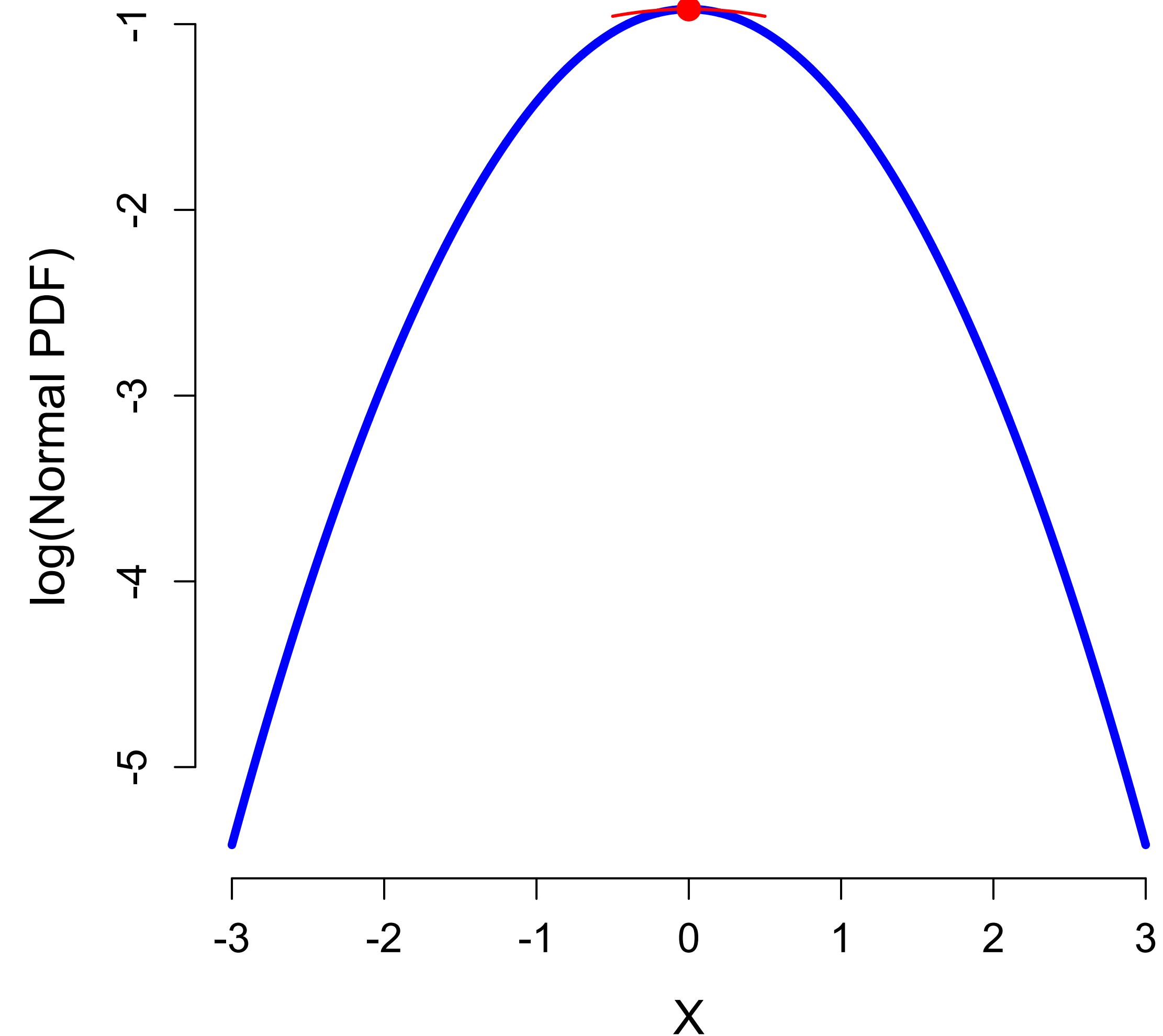
LAPLACE APPROXIMATION

- ▶ LA first finds the maximum of the log posterior
(maximum a posteriori inference)



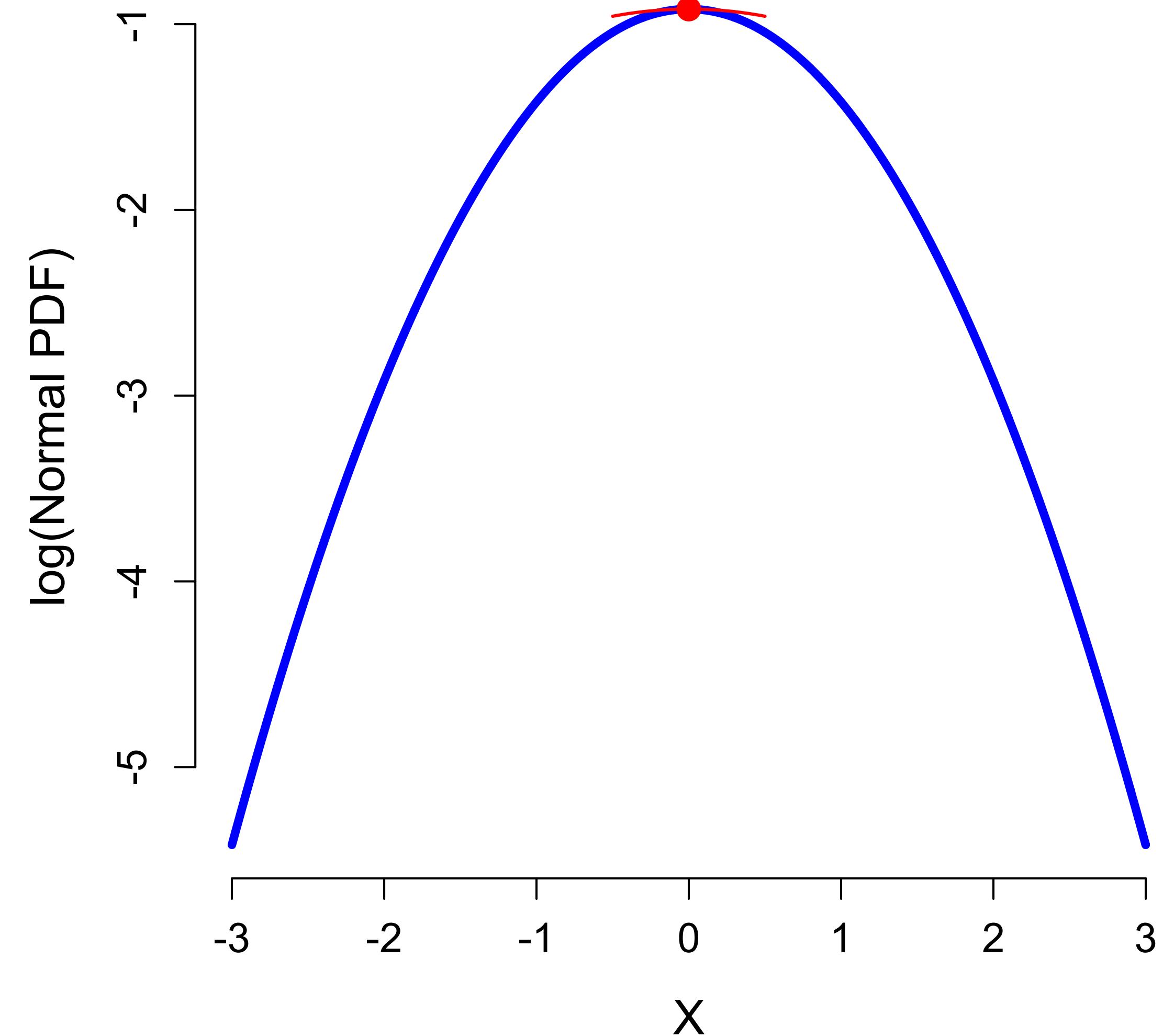
LAPLACE APPROXIMATION

- ▶ LA first finds the maximum of the log posterior (maximum a posteriori inference)
- ▶ We then estimate the **hessian matrix** to approximate the curvature at the maximum



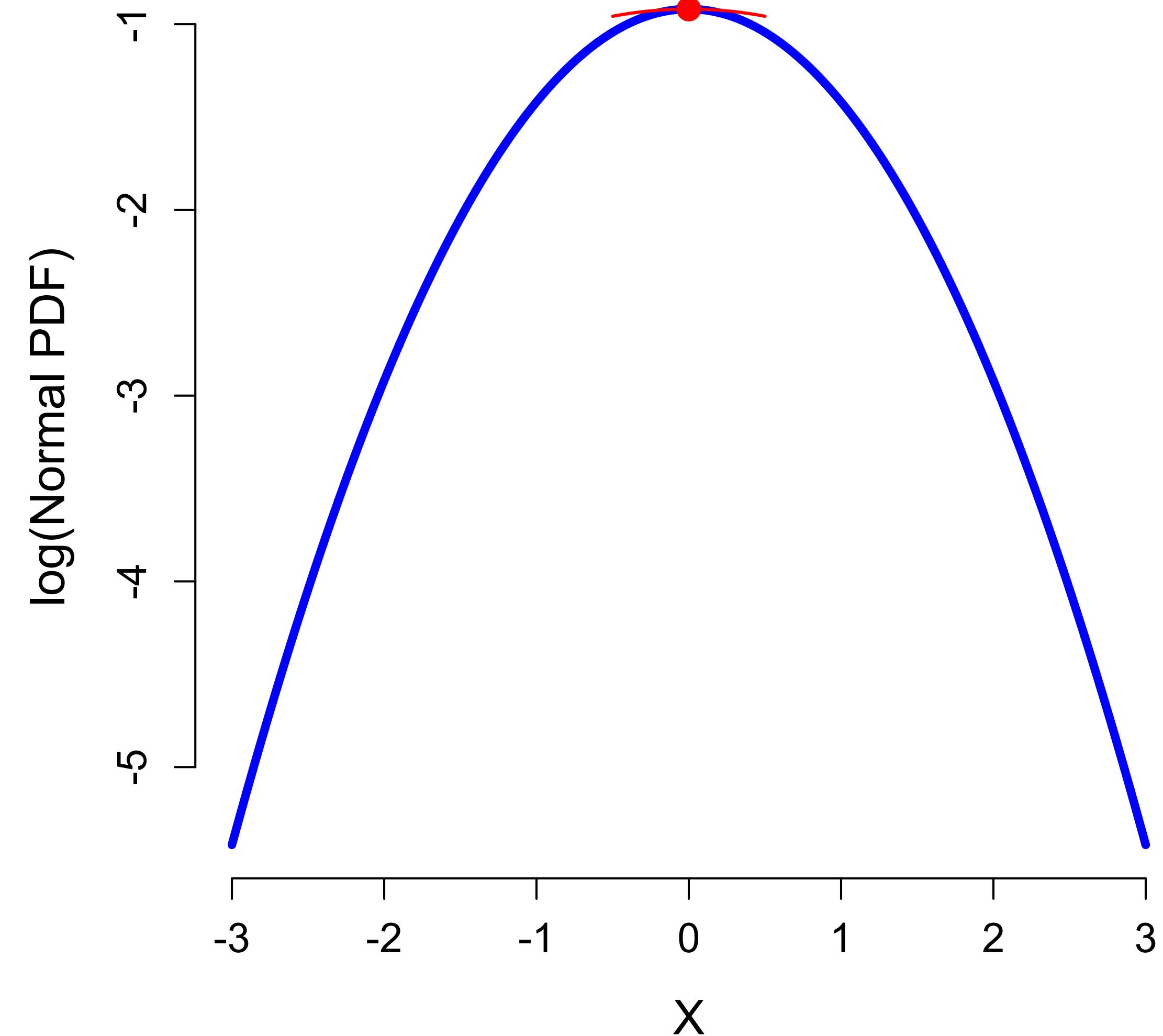
LAPLACE APPROXIMATION

- ▶ LA first finds the maximum of the log posterior (maximum a posteriori inference)
- ▶ We then estimate the **hessian matrix** to approximate the curvature at the maximum
- ▶ The hessian can be solved to estimate a **variance-covariance matrix** for all parameters



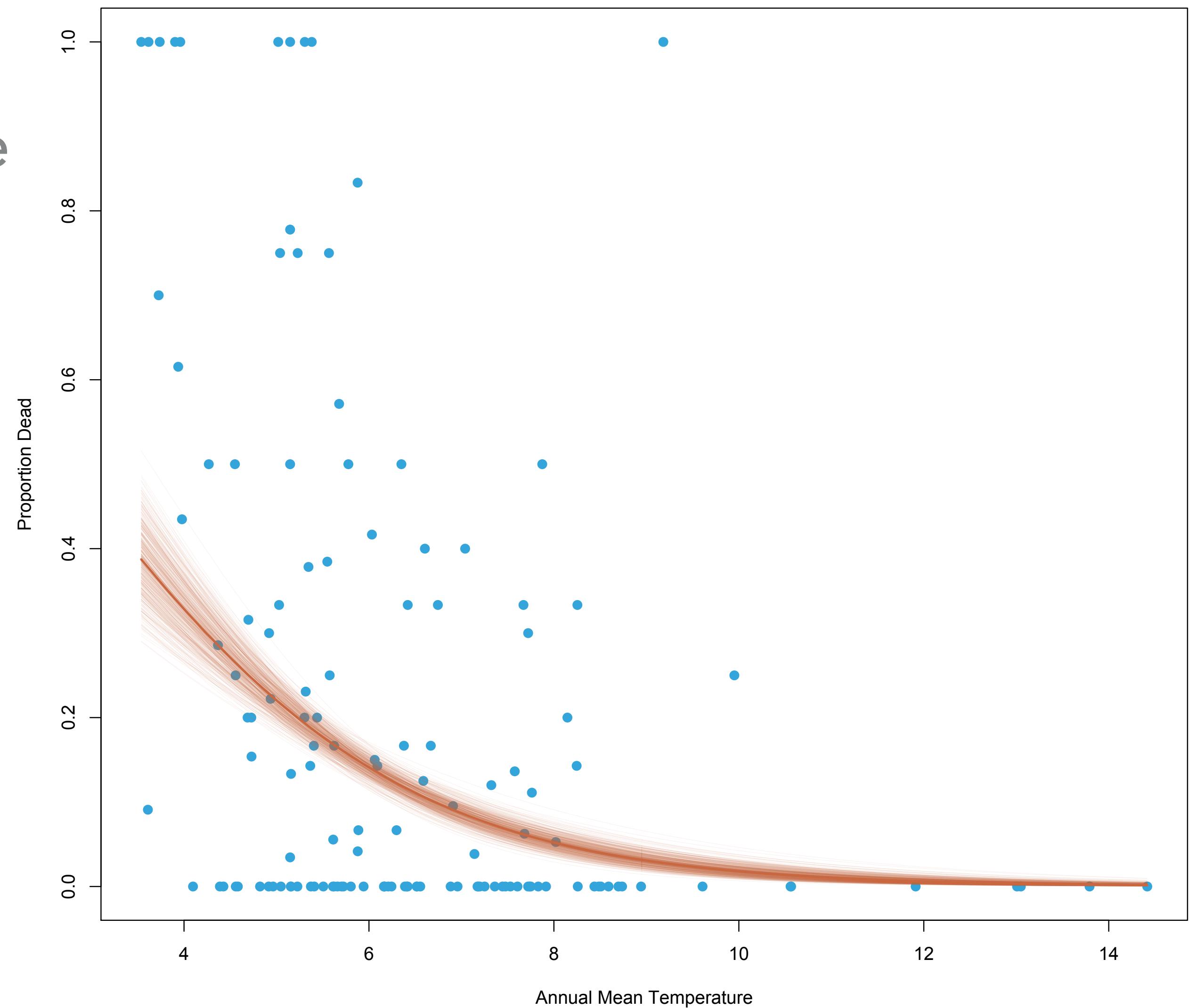
LAPLACE APPROXIMATION

- ▶ LA first finds the maximum of the log posterior (maximum a posteriori inference)
- ▶ We then estimate the **hessian matrix** to approximate the curvature at the maximum
- ▶ The hessian can be solved to estimate a **variance-covariance matrix** for all parameters
- ▶ Finally, we draw samples from a multivariate normal distribution to make posterior inference



GLM COMPARISON

- ▶ Using Stan, we found a clear negative relationship between mortality and temperature.
- ▶ How do the results compare using Laplace approximation?



BINOMIAL GLM

- ▶ This model has two parameters: a slope and an intercept

```
log_posterior <- function(params, data) {  
  ntrees <- data$n # number of trees in each plot  
  died <- data$died # number of trees that died  
  temperature <- data$annual_mean_temp  
  
  intercept <- params[1]  
  slope <- params[2]  
  
  # using a logit link function  
  mu <- plogis(intercept + slope * temperature)  
  
  ll <- sum(dbinom(died, ntrees, mu, log=TRUE))  
  lp <- dnorm(intercept, 0, 10, log = TRUE) +  
    dnorm(slope, 0, 5, log=TRUE)  
  
  return(ll + lp)  
}
```

BINOMIAL GLM

- ▶ Then it is simply a matter of using optim with hessian=TRUE
- ▶ MAP estimate & standard errors is nearly identical to the means from Stan
- ▶ Solve the hessian matrix to get the standard error
- ▶ Can use rmvnorm to get posterior samples
- ▶ Using Laplace approximation is 240-3300x faster

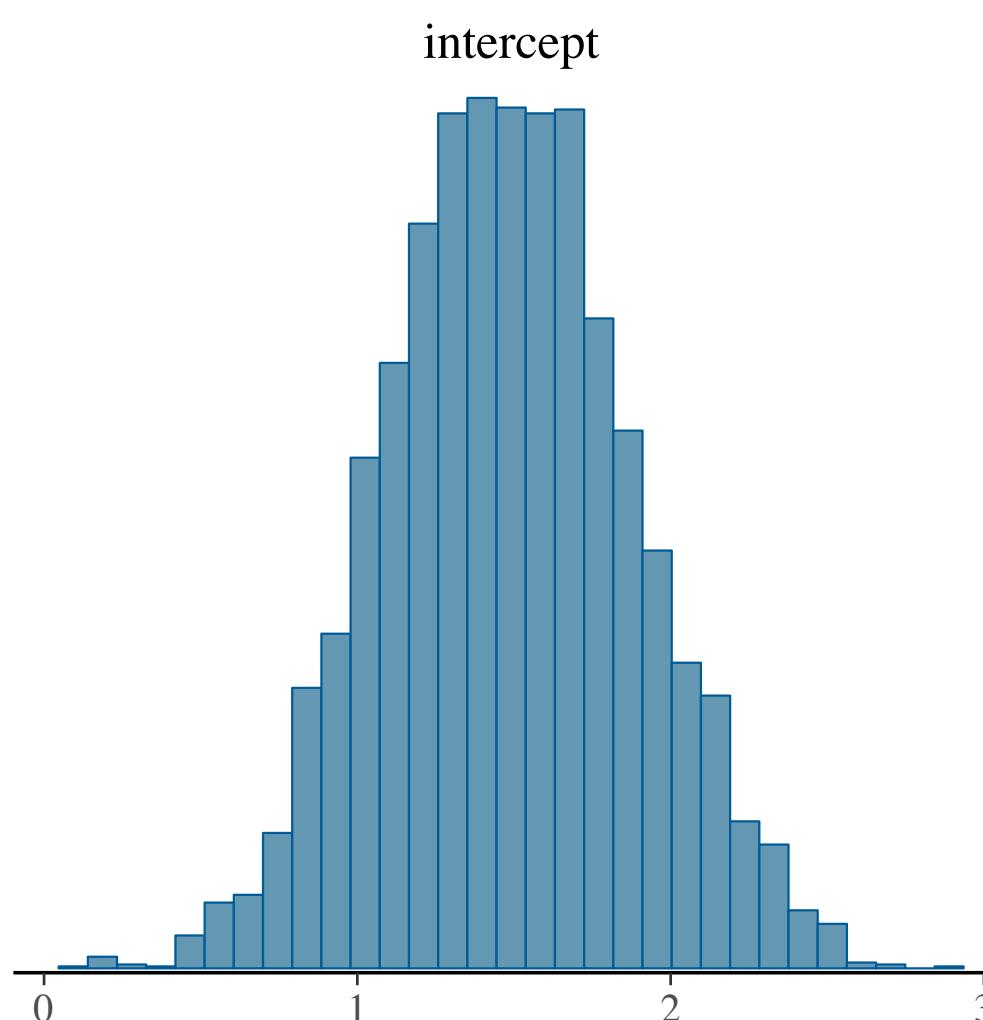
```
inits <- c(rnorm(1, 0, 10), rnorm(1, 0, 5))

fit <- optim(inits, log_posterior, control =
  list(fnscale = -1), hessian = TRUE, data = dat)
fit$par
vcv_mat <- solve(-fit$hessian)
## the standard error of each param
sqrt(diag(vcv_mat))
```

	Stan	Laplace
intercept	1.484 (0.39)	1.478 (0.40)
slope	-0.549 (0.069)	-0.547 (0.071)
Time to fit (s)	2.407 (33.192)	0.010

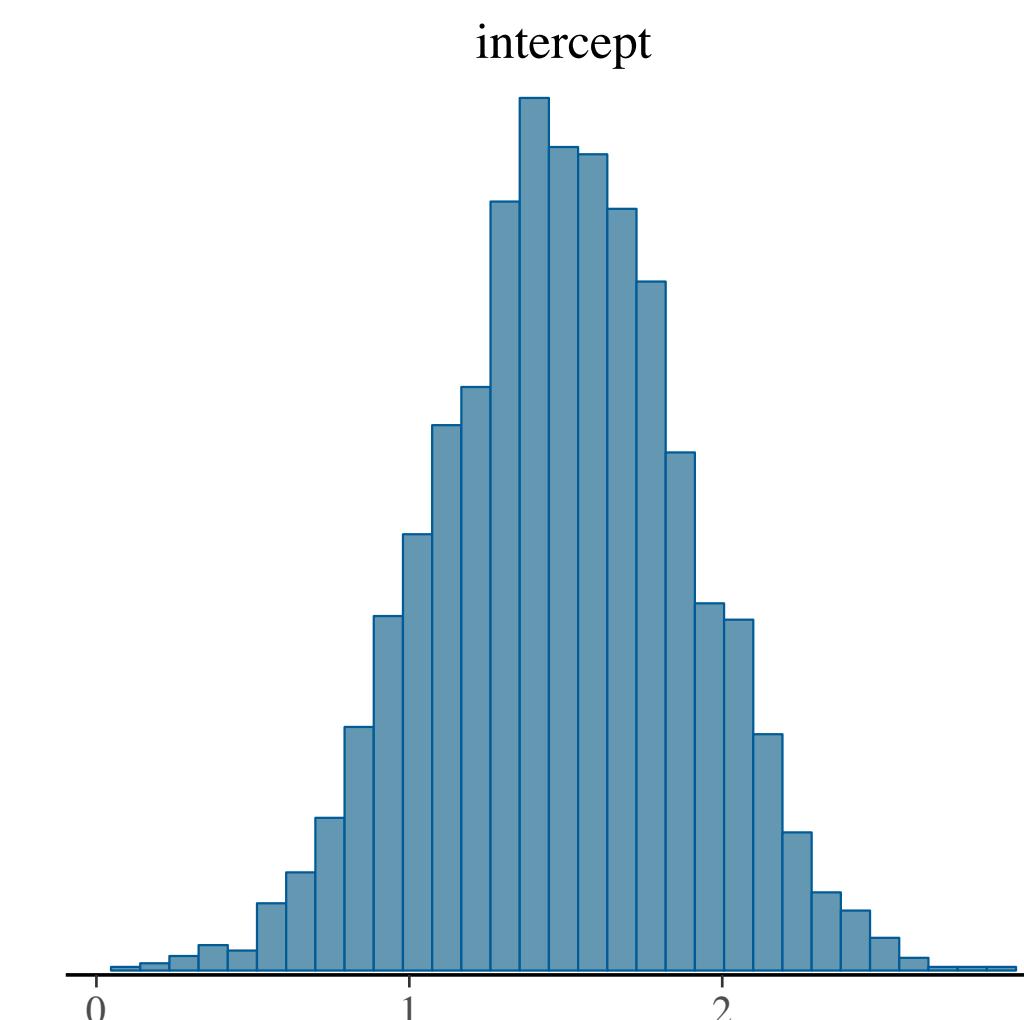
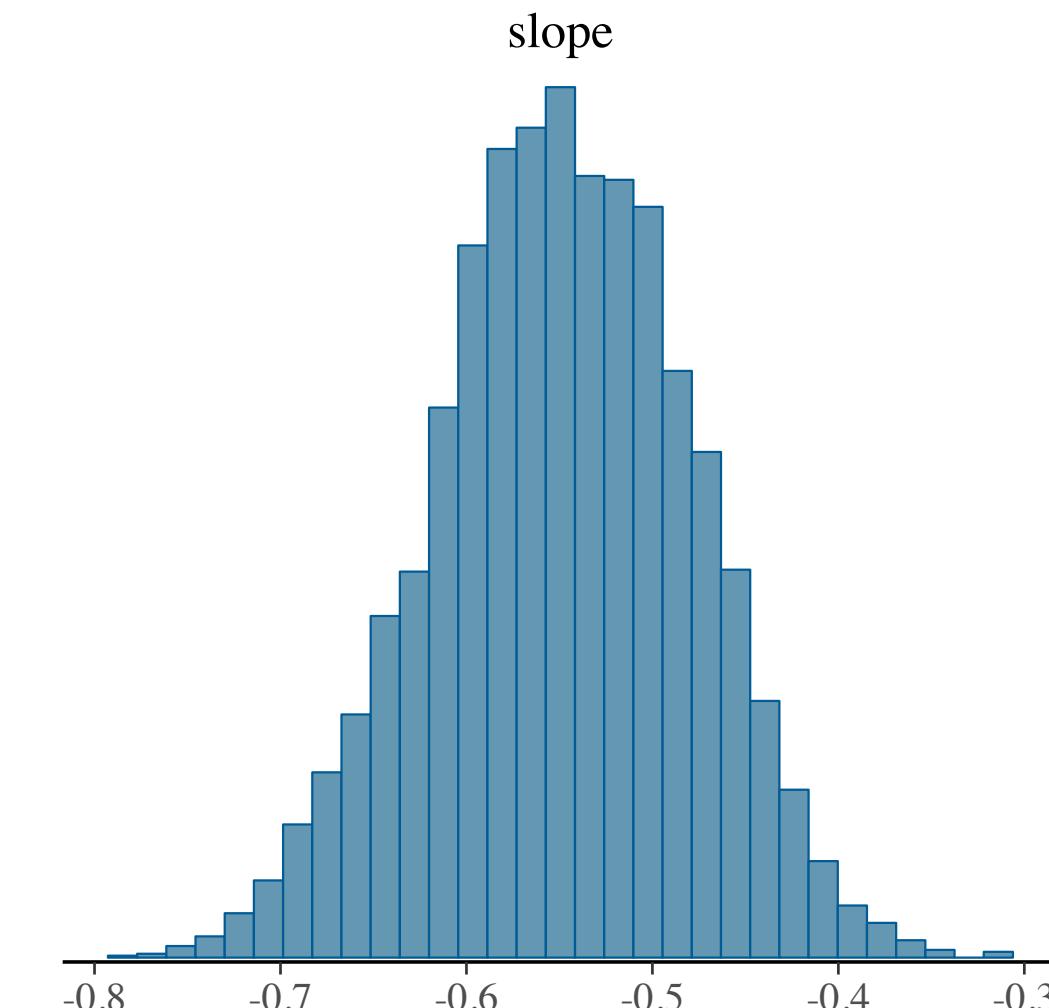
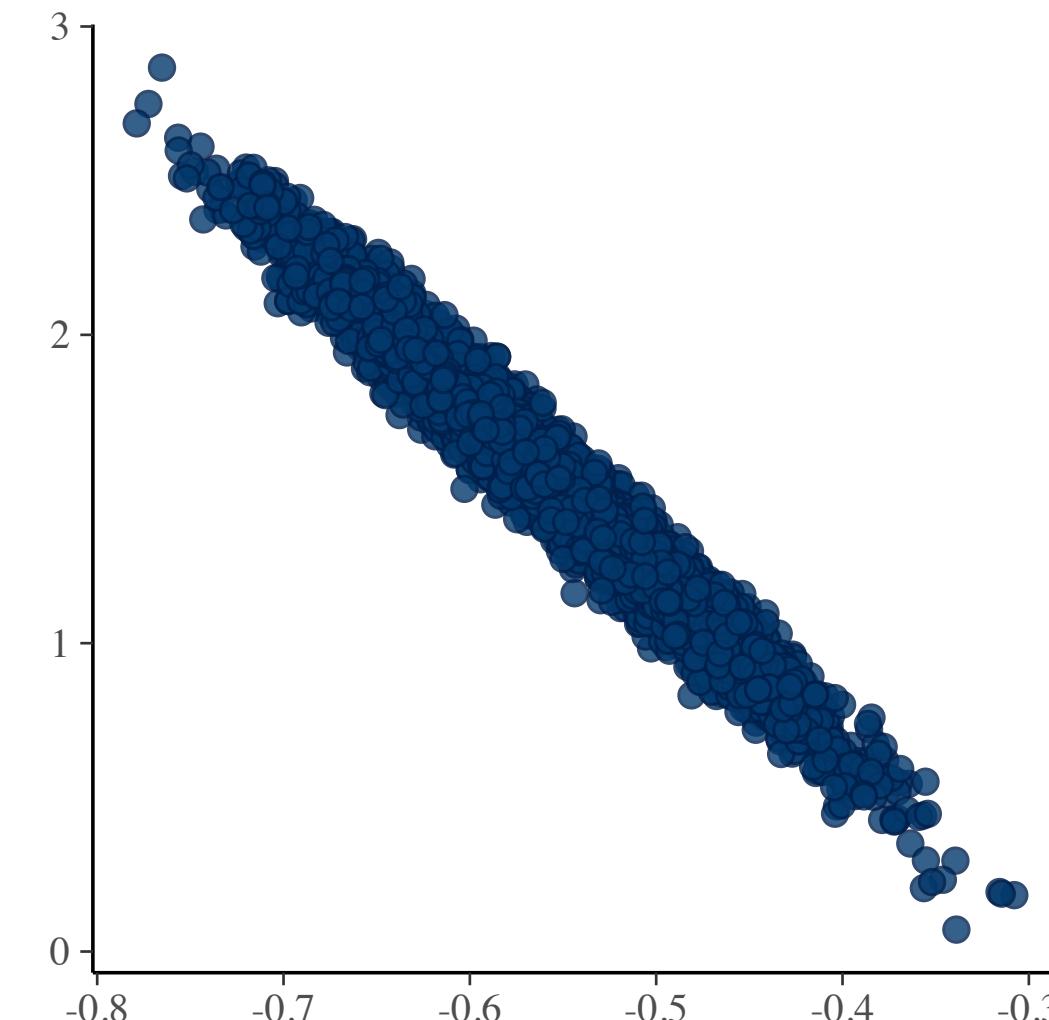
LAPLACE APPROXIMATION

GLM PAIRS PLOTS



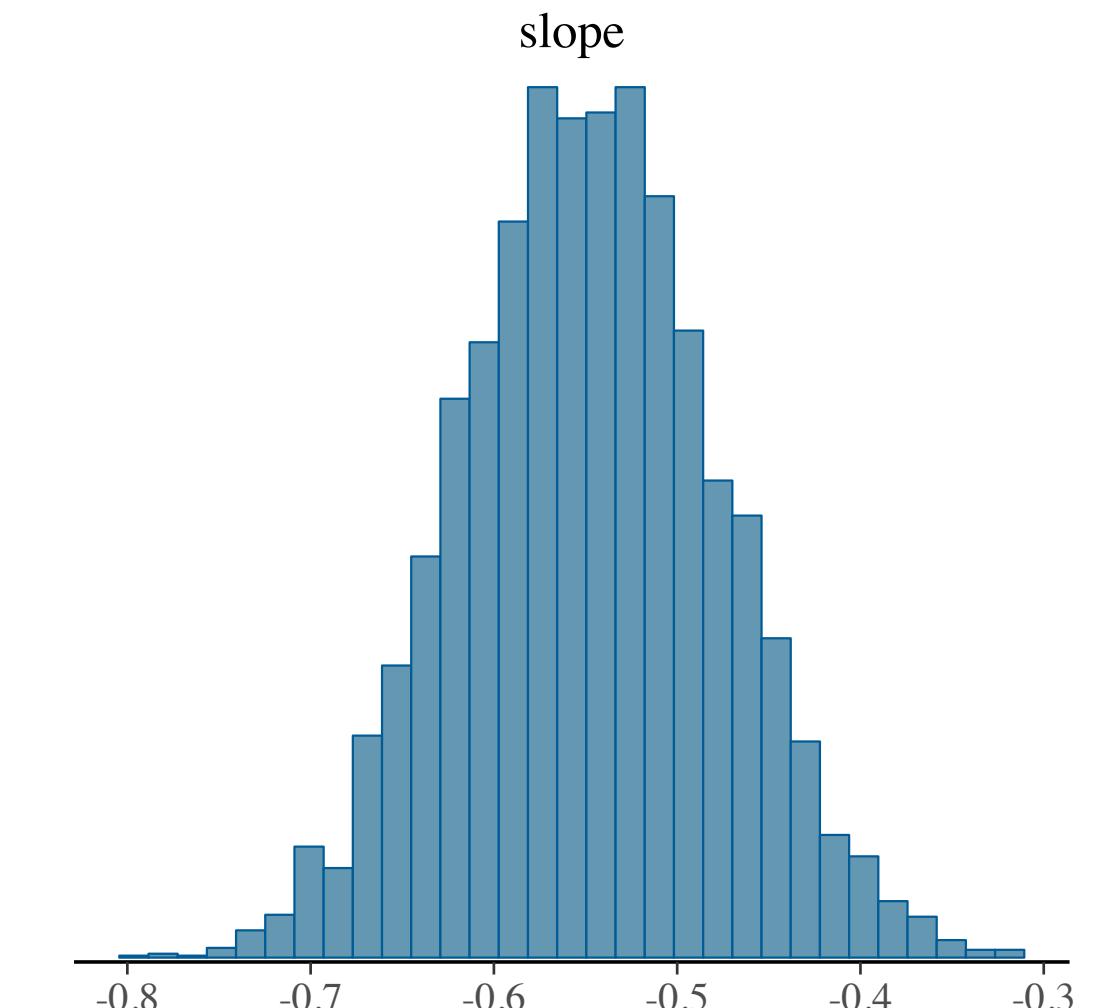
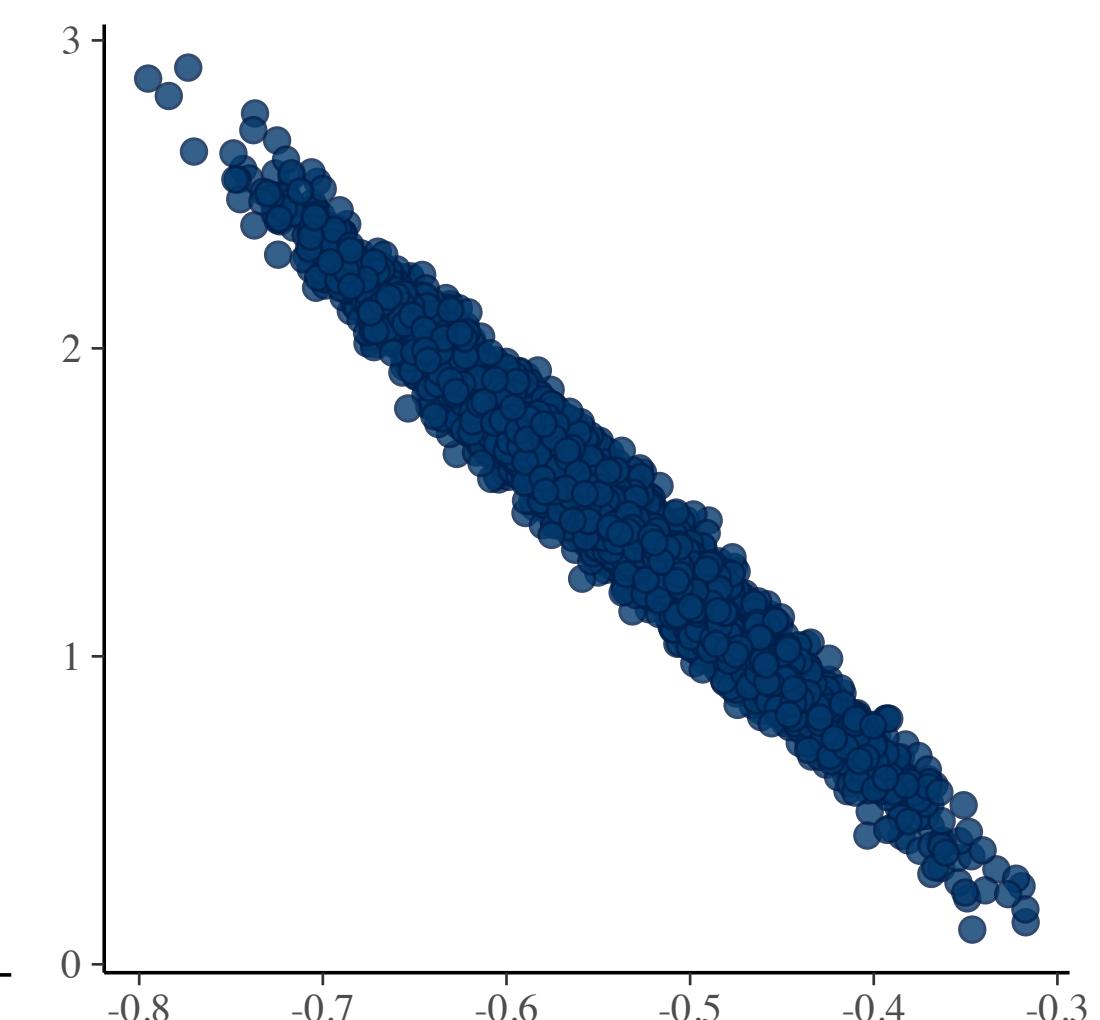
Stan

$r = -0.978$



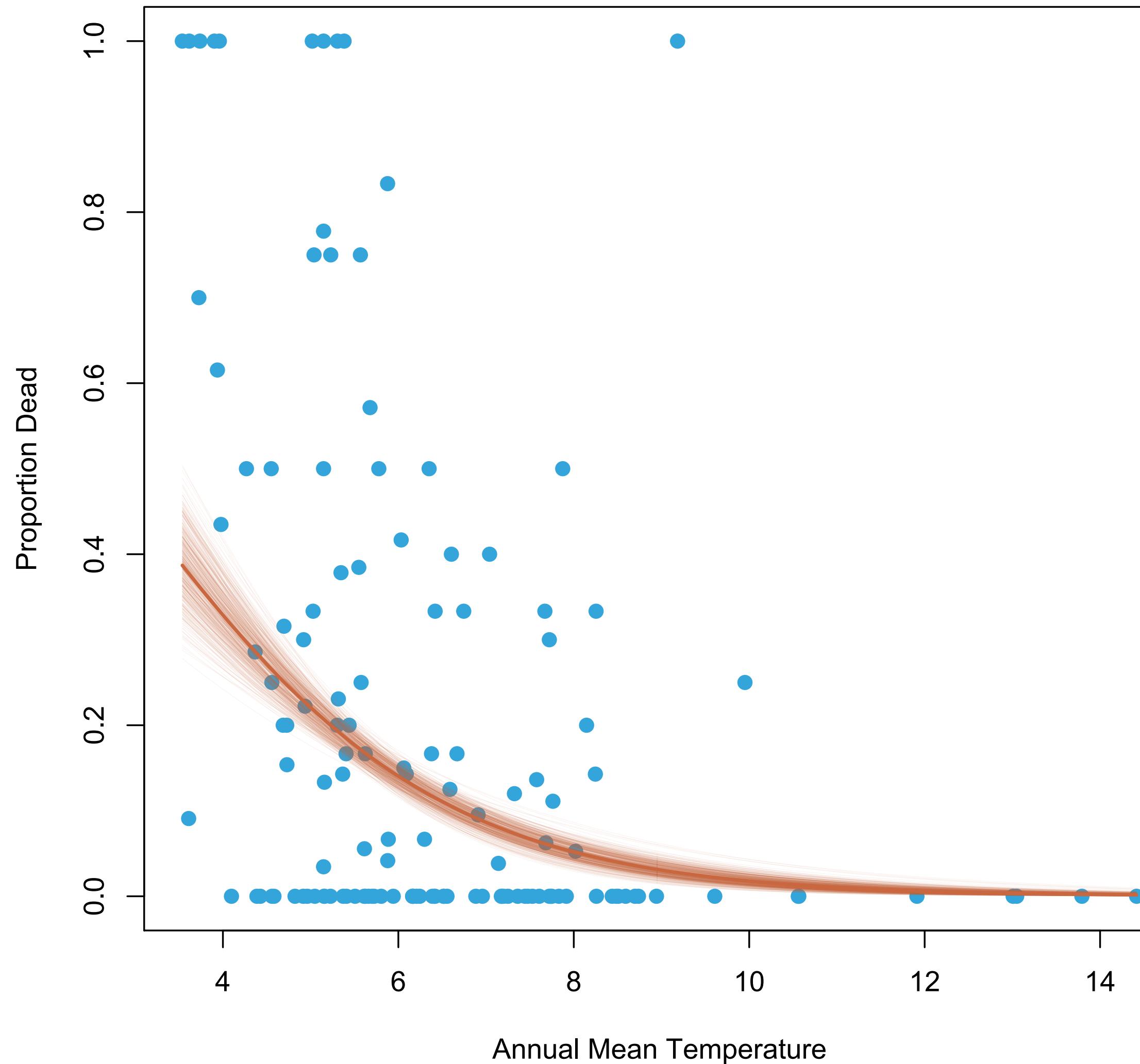
Laplace

$r = -0.979$



PREDICTIONS COMPARISON

Stan



Laplace

