Let $y_{1j}$ be the time to abortion of the $j$th cow infected with *Neospora caninum* and let $y_{2j}$ be the time to abortion of the $j$th cow in the uninfected group. We assume that the logs of the abortion times in each group are normal with means $\mu_1$, $\mu_2$ and variances $\sigma_1^2$, $\sigma_2^2$. Referring to the abortion times as having *log-normal distributions*, write

$$y_{11},\ldots,y_{1,19}|\mu_1,\sigma_1^2 \overset{iid}{\sim} LN(\mu_1,\sigma_1^2)$$

and

$$y_{21},\ldots,y_{2,26}|\mu_2,\sigma_2^2 \overset{iid}{\sim} LN(\mu_2,\sigma_2^2).$$

We use independent reference priors

$$\mu_i \sim N(0,1000) \quad \text{and} \quad 1/\sigma_i^2 \sim \text{Gamma}(0.001,0.001), \quad i=1,2.$$

Given the model parameters, the median times to abortion in the two groups are $\exp(\mu_1)$ and $\exp(\mu_2)$ so the difference in medians is $\Delta = \exp(\mu_1) - \exp(\mu_2)$. We are interested in the posterior distribution $\Delta|y_{11},\ldots,y_{1,19},y_{21},\ldots,y_{2,26}$ and whether the difference is reasonably close to zero.

Computer simulations were used to estimate the posterior median of $\Delta$, which is 27.8 days, and the equal-tailed 95% probability interval for $\Delta$, which is (1.7, 55.2) days. With 95% probability, infected cows have a median time to abortion between 1.7 and 55.2 days longer than those cows not infected with *Neospora caninum*. If we use alternative reference priors that are diffuse but finite uniform distributions on the means and precisions, these numbers change only slightly; the posterior median is 28.0 days and the 95% probability interval is 1.8 to 54.9 days.

## 1.5 Ache Hunting with Age Trends

We now return to Ache hunting with more data and a more sophisticated model that incorporates random hunter effects into the analysis along with a tendency for hunting success to change with age. Data were collected on the daily number of armadillos killed by 38 adult males of an Ache tribe over several forest treks. There are 1,302 total observations $y_{ij}$ where $i = 1,\ldots,38$ indexes an Ache male and $j = 1,\ldots,n_i$ denotes a day spent hunting. A plot of the average number of kills per man by age shows a generally increasing, then decreasing trend. It is of interest to model and quantify how a man's age affects daily kill success. We assume the number of armadillos killed is distributed Poisson and take the log-rate to be a quadratic function of a man's age in years $a_i$ plus a subject-specific, normally distributed random effect $\delta_i$. The random effect for each man accounts for the correlation of an individual's daily kills over the several hunting trips in which the data were collected. One might also think of $\delta_i$ as the innate ability of hunter $i$.

The sampling model is specified

$$\begin{aligned} y_{ij}|\lambda_i &\overset{ind}{\sim} \text{Pois}(\lambda_i), \quad i=1,\ldots,38; j=1,\ldots,n_i, \\ \log(\lambda_i) &= \beta_1 + \beta_2(a_i-\bar{a}) + \beta_3(a_i-\bar{a})^2 + \delta_i \\ \delta_i|\tau &\overset{iid}{\sim} N(0,\tau^{-1}). \end{aligned}$$

Here $\lambda_i$ is the mean daily kill rate for individual $i$, $\bar{a}$ is the average age of the 38 hunters, while the variance of the normal distribution, usually denoted $\sigma^2$, has been reparameterized as the *precision*

$$\tau \equiv \frac{1}{\sigma^2},$$

which turns out to be much more convenient for Bayesian analysis of normal distributions. The ages $a_i$ are fixed, known constants. The parameters $\beta_1$, $\beta_2$, $\beta_3$, and $\tau$ are given independent reference prior distributions $\beta_i \sim N(0,1000)$, $i = 1,2,3$ and $\tau \sim \text{Gamma}(0.001,0.001)$.

Table 1.1: *Armadillo kills: posterior medians and probability intervals.*

| Parameter | Median | 95% PI |
|---|---|---|
| $\beta_1$ | $-0.7147$ | $(-1.007, -0.433)$ |
| $\beta_2$ | 0.01368 | (-0.002525, 0.03085) |
| $\beta_3$ | $-0.002683$ | $(-0.004007, -0.001459)$ |
| $\sigma$ | 0.4252 | (0.2731, 0.658) |

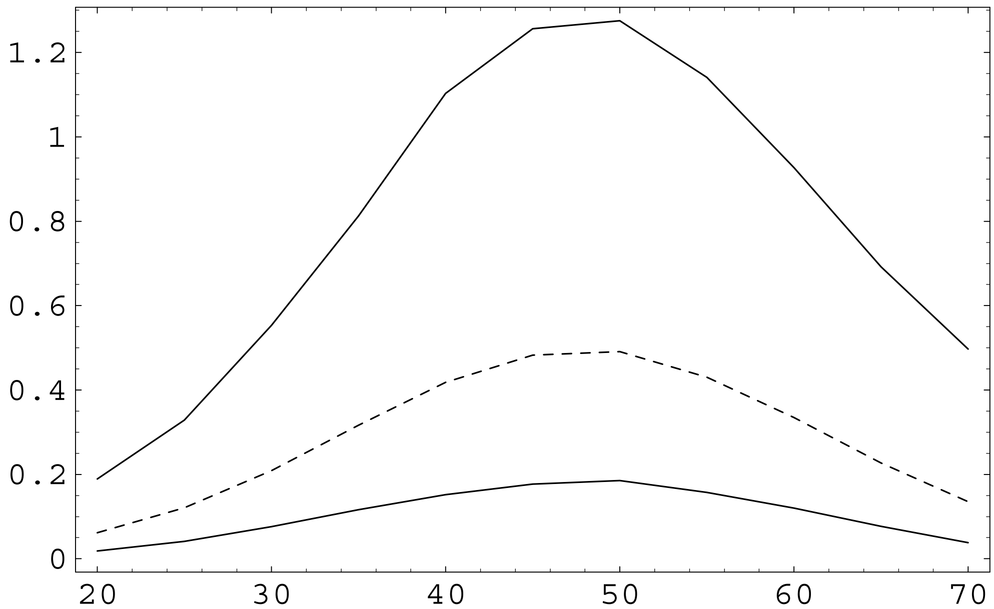


Figure 1.2: *Estimated mean daily kill by age with 95% PI.*

We obtain posterior information by simulation. Summary information is presented in Table 1.1. The quadratic coefficient $\beta_3$ is estimated to be $-0.0027$ and from the probability interval it is clearly nonzero. If we exponentiate $\log(\lambda)$, we obtain an estimate of the mean daily kill rate as a function of age $a$,

$$\hat{\lambda}(a) = \exp\{-0.7147 + 0.01368(a - \bar{a}) - 0.002683(a - \bar{a})^2\},$$

although this estimate is not necessarily the posterior median.

Posterior medians for the mean daily kill $\lambda(a)$ and probability intervals were computed for ages 20 to 70 in steps of five years to obtain Figure 1.2. The range of ages in the data is 20 to 66 years. We see that the average kill-rate increases with age up until about 50, perhaps reflecting that hunting experience increases the chance of killing an armadillo, but then declines as the hunter enters his "golden years." When the model was refit using independent, infinite, and improper priors on all model parameters the resulting posterior inferences were almost identical to those presented above.

## 1.6 Lung Cancer Treatment: Log-Normal Regression

The drugs cisplaten and etoposide can increase the lifetimes of those with limited-stage small cell lung cancer. It was of interest to determine which sequencing of the drugs works better. Treatment 0 was the administration of cisplaten followed by etoposide. Those receiving treatment 1 were given etoposide followed by cisplaten. The 121 patients studied were randomly assigned to the two treatment groups: 62 patients received treatment 0 and 59 patients received treatment 1. The data are the time in days $y$ that a patient was known to be alive from the start of the treatment regimen, along