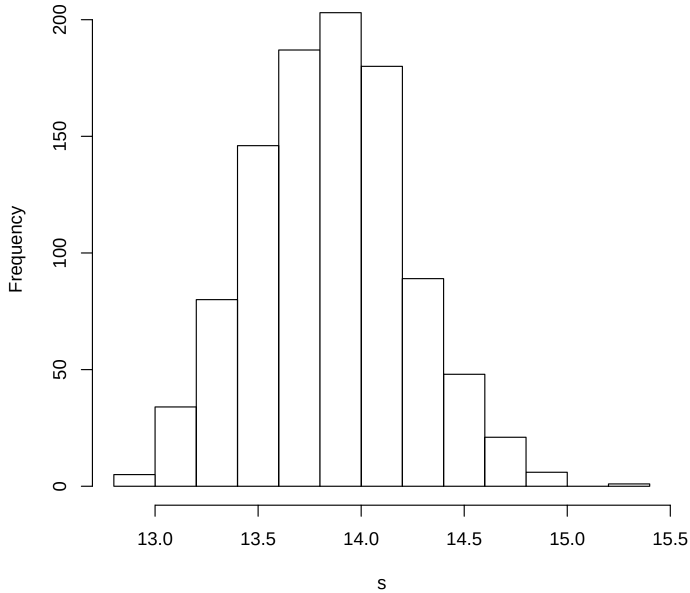


**Fig. 3.1.** Histogram of simulated sample of the standard deviation $\sigma$ of differences between game outcomes and point spreads.

## 3.3 Estimating a Heart Transplant Mortality Rate

Consider the problem of learning about the rate of success of heart transplant surgery of a particular hospital in the United States. For this hospital, we observe the number of transplant surgeries $n$, and the number of deaths within 30 days of surgery $y$ is recorded. In addition, one can predict the probability of death for an individual patient. This prediction is based on a model that uses information such as patients' medical condition before surgery, gender, and race. Based on these predicted probabilities, one can obtain an expected number of deaths, denoted by $e$. A standard model assumes that the number of deaths $y$ follows a Poisson distribution with mean $e\lambda$, and the objective is to estimate the mortality rate per unit exposure $\lambda$.

The standard estimate of $\lambda$ is the maximum likelihood estimate $\hat{\lambda} = y/e$. Unfortunately, this estimate can be poor when the number of deaths $y$ is close to zero. In this situation when small death counts are possible, it is desirable to use a Bayesian estimate that uses prior knowledge about the size of the mortality rate. A convenient choice for a prior distribution is a member of the gamma$(\alpha, \beta)$ density of the form

$$p(\lambda) \propto \lambda^{\alpha-1} \exp(-\beta\lambda), \ \lambda > 0.$$

A convenient source of prior information is heart transplant data from a small group of hospitals that we believe has the same rate of mortality as the rate from the hospital of interest. Suppose we observe the number of deaths $z_j$ and the exposure $o_j$ for ten hospitals ($j = 1, ..., 10$), where $z_j$ is Poisson with mean $o_j\lambda$. If we assign $\lambda$ the standard noninformative prior $p(\lambda) \propto \lambda^{-1}$, then the updated distribution for $\lambda$, given these data from the ten hospitals, is

$$p(\lambda) \propto \lambda^{\sum_{j=1}^{10} z_j - 1} \exp\left(-(\sum_{j=1}^{10} o_j)\lambda\right).$$

Using this information, we have a gamma$(\alpha, \beta)$ prior for $\lambda$, where $\alpha = \sum_{j=1}^{10} z_j$ and $\beta = \sum_{j=1}^{10} o_j$. In this example, we have

$$\sum_{j=1}^{10} z_j = 16, \sum_{j=1}^{10} o_j = 15174,$$

and so we assign $\lambda$ a gamma(16, 15174) prior.

If the observed number of deaths from surgery $y_{\text{obs}}$ for a given hospital with exposure $e$ is Poisson $(e\lambda)$ and $\lambda$ is assigned the gamma$(\alpha, \beta)$ prior, then the posterior distribution will also have the gamma form with parameters $\alpha + y_{\text{obs}}$ and $\beta + e$. Also the (prior) predictive density of $y$ (before any data are observed) can be computed using the formula

$$f(y) = \frac{f(y|\lambda)g(\lambda)}{g(\lambda|y)},$$

where $f(y|\lambda)$ is the Poisson$(e\lambda)$ sampling density and $g(\lambda)$ and $g(\lambda|y)$ are, respectively, the prior and posterior densities of $\lambda$.

By the model-checking strategy of Box (1980), both the posterior density $g(\lambda|y)$ and the predictive density $f(y)$ play important roles in a Bayesian analysis. By using the posterior density, one performs inference about the unknown parameter conditional on the Bayesian model that includes the assumptions of sampling density and the prior density. One can check the validity of the proposed model by inspecting the predictive density. If the observed data value $y_{\text{obs}}$ is consistent with the predictive density $p(y)$, then the model seems reasonable. On the other hand, if $y_{\text{obs}}$ is in the extreme tail portion of the predictive density, then this casts doubt on the validity of the Bayesian model, and perhaps the prior density or the sampling density has been misspecified.

We consider inference about the heart transplant death rate for two hospitals – one that has experienced a small number of surgeries and a second that has experienced many surgeries. First consider hospital A, which experienced only one death ($y_{\text{obs}} = 1$) with an exposure of $e = 66$. The standard estimate of this hospital's rate, 1/66, is suspect due to the small observed number of deaths. The following R calculations illustrate the Bayesian calculations. After the gamma prior parameters `alpha` and `beta` and exposure `ex` are defined, the predictive density of the values $y = 0, 1, ..., 10$ is found by using the preceding formula and the R functions `dpois` and `dgamma`. The formula for the predictive density is valid for all $\lambda$, but to ensure that there is no underflow in the calculations, the values of $f(y)$ are computed for the prior mean value $\lambda = \alpha/\beta$. Note that practically all of the probability of the predictive density is concentrated on the two values $y = 0$ and 1. The observed number of deaths ($y_{\text{obs}} = 1$) is in the middle of this predictive distribution, so there is no reason to doubt our Bayesian model.

```
> alpha=16;beta=15174
> yobs=1; ex=66
> y=0:10
> lam=alpha/beta
> py=dpois(y, lam*ex)*dgamma(lam, shape = alpha,
+   rate = beta)/dgamma(lam, shape= alpha + y,
+   rate = beta + ex)
> cbind(y, round(py, 3))
```

```
       y
 [1,]  0 0.933
 [2,]  1 0.065
 [3,]  2 0.002
 [4,]  3 0.000
 [5,]  4 0.000
 [6,]  5 0.000
 [7,]  6 0.000
 [8,]  7 0.000
 [9,]  8 0.000
[10,]  9 0.000
[11,] 10 0.000
```

The posterior density of $\lambda$ can be summarized by simulating 1000 values from the gamma density.

```
> lambdaA = rgamma(1000, shape = alpha + yobs, rate = beta + ex)
```

Let's consider the estimation of a different hospital that experiences many surgeries. Hospital B had $y_{\text{obs}} = 4$ deaths, with an exposure of $e = 1767$. For these data, we again have R compute the prior predictive density and simulate 1000 draws from the posterior density using the `rgamma` command. Again we see that the observed number of deaths seems consistent with this model since $y_{\text{obs}} = 4$ is not in the extreme tails of this distribution.

```
> ex = 1767; yobs=4
> y = 0:10
```

```
> py = dpois(y, lam * ex) * dgamma(lam, shape = alpha,
+     rate = beta)/dgamma(lam, shape = alpha + y,
+     rate = beta + ex)
> cbind(y, round(py, 3))

       y
 [1,]  0 0.172
 [2,]  1 0.286
 [3,]  2 0.254
 [4,]  3 0.159
 [5,]  4 0.079
 [6,]  5 0.033
 [7,]  6 0.012
 [8,]  7 0.004
 [9,]  8 0.001
[10,]  9 0.000
[11,] 10 0.000

> lambdaB = rgamma(1000, shape = alpha + yobs, rate = beta + ex)
```

To see the impact of the prior density on the inference, it is helpful to display the prior and posterior distributions on the same graph. In Figure 3.2, density estimates of the simulated draws from the posterior distributions of the rates are shown for hospitals A and B. The gamma prior density is also displayed in each case. We see that for hospital A, with relatively little experience in surgeries, the prior information is significant and the posterior distribution resembles the prior distribution. In contrast, for hospital B, with many surgeries, the prior information is less influential and the posterior distribution resembles the likelihood function.

```
> par(mfrow = c(2, 1))
> plot(density(lambdaA), main="HOSPITAL A",
+   xlab="lambdaA", lwd=3)
> curve(dgamma(x, shape = alpha, rate = beta), add=TRUE)
> legend("topright",legend=c("prior","posterior"),lwd=c(1,3))
> plot(density(lambdaB), main="HOSPITAL B",
+   xlab="lambdaB", lwd=3)
> curve(dgamma(x, shape = alpha, rate = beta), add=TRUE)
> legend("topright",legend=c("prior","posterior"),lwd=c(1,3))
```

## 3.4 An Illustration of Bayesian Robustness

In practice, one may have incomplete prior information about a parameter in the sense that one's beliefs won't entirely define a prior density. There may be a number of different priors that match the given prior information. For