Chapter 8.5 Updating the Normal Prior

Jim Albert and Monika Hu

Chapter 8 Modeling Measurement and Count Data

Introduction

- Current prior beliefs about Federer's mean time-to-serve μ are represented by a Normal curve with mean 18 seconds and standard deviation 1.56 seconds.
- Data is collected — Federer's time-to-serves are recorded for 20 serves and the sample mean is 17.2 seconds.

Likelihood

- With substitution of the values $\bar{y} = 17.2$, $n = 20$, and $\sigma = 4$, we obtain

$$L(\mu) \propto \exp\left\{-\frac{20}{2(4)^2}(17.2-\mu)^2\right\}$$

- Viewing the likelihood as a function of the parameter μ , the likelihood is recognized as a Normal density with mean $\bar{y} = 17.2$ and standard deviation $\sigma/\sqrt{n} = 4/\sqrt{20} = 0.89$.

Conjugate analysis

- One obtains the posterior density curve by multiplying the Normal prior by the likelihood.
- Working through some algebra, one will see that the posterior density also has the Normal density form.
- The Normal prior is said to be *conjugate* since the prior and posterior densities come from the same distribution family: Normal.
- To be specific, if one has a Normal prior for the unknown mean μ with mean μ_0 and standard deviation σ_0, one obtains a Normal posterior for μ with updated parameters μ_n and σ_n.

Summary of the update procedure

- In table, there are rows corresponding to Prior, Data/Likelihood, and Posterior and columns corresponding to Mean, Precision, and Standard Deviation.
- We put in the information we know.

Type	Mean	Precision	Stand
Prior	18.00		1.56
Data/Likelihood	17.20		0.89
Posterior			

Precision

- We define the *precision*, ϕ, to be the reciprocal of the square of the standard deviation.
- We compute the precisions of the prior and data from the given standard deviations and fill in the Precision column of the table.

Type	Mean	Precision	Stand
Prior	18.00	0.41	1.56
Data/Likelihood	17.20	1.26	0.89
Posterior			

Precisions add

- The Posterior precision is the sum of the Prior precision and the Data/Likelihood precisions:

$$\phi_{post} = \phi_{prior} + \phi_{data} = 0.41 + 1.26 = 1.67.$$

- The posterior standard deviation is computed as the reciprocal of the square root of the precision.

$$\sigma_n = \frac{1}{\sqrt{\phi_{post}}} = \frac{1}{\sqrt{1.67}} = 0.77.$$

Adding to table

- These standard deviations are entered into the table.

Type	Mean	Precision	Stand
Prior	18.00	0.41	1.56
Data/Likelihood	17.20	1.26	0.89
Posterior		1.67	0.77

Posterior mean

- The posterior mean is a weighted average of the Prior and Data/Likelihood means where the weights are given by the corresponding precisions.
- That is, the formula is given by

$$\mu_n = \frac{\phi_{prior} \times \mu_0 + \phi_{data} \times \bar{y}}{\phi_{prior} + \phi_{data}}.$$

- By making appropriate substitutions, we obtain the posterior mean:

$$\mu_n = \frac{0.41 \times 18.00 + 1.26 \times 17.20}{0.41 + 1.26} = 17.40.$$

The posterior

The posterior density is Normal with mean 17.40 seconds and standard deviation 0.77 seconds.

Type	Mean	Precision	Stand
Prior	18.00	0.41	1.56
Data/Likelihood	17.20	1.26	0.89
Posterior	17.40	1.67	0.77

Using R

The Normal updating is performed by the R function `normal_update()`. One inputs two vectors – `prior` is a vector of the prior mean and standard deviation and `data` is a vector of the sample mean and standard error. The output is a vector of the posterior mean and posterior standard deviation.

```
prior <- c(18, 1.56)
data <- c(17.20, 0.89)
normal_update(prior, data)
```

```
## [1] 17.3964473  0.7730412
```

Plotting prior and posterior

- The posterior has a smaller spread since the posterior has more information than the prior about Federer's mean time-to-serve.

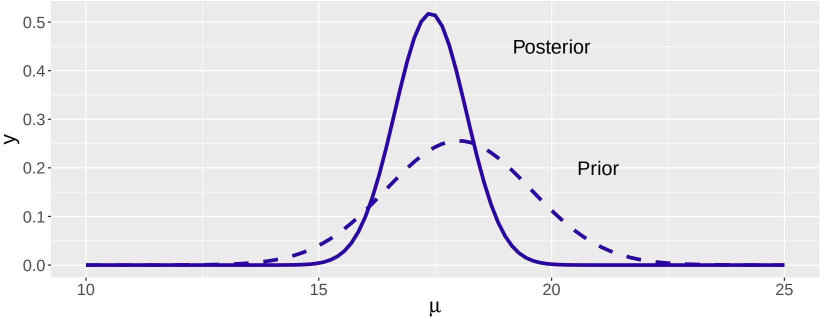

Figure 1: Prior and posterior curves for Federer's mean time-to-serve μ.

Summary of our work

- We collect a sample of data $Y_1, \ldots, Y_n$ that is $\mathrm{Normal}(\mu, \sigma)$ where we assume the sampling standard deviation σ is known.
- We assign a $\mathrm{Normal}(\mu_0, \sigma_0)$ prior to the mean μ
- After $Y_1 = y_1, \ldots, Y_n = y_n$ are observed, the posterior distribution for the mean μ is another Normal distribution
- Posterior mean is $\frac{\phi_0 \mu_0 + n\phi \bar{y}}{\phi_0 + n\phi}$
- Posterior precision is $\phi_0 + n\phi$ (thus standard deviation $\sqrt{\frac{1}{\phi_0 + n\phi}}$):

Posterior compromises between the prior and the sample

- One can rewrite the posterior mean as

$$\mu_n = \frac{\phi_0}{\phi_0 + n\phi}\mu_0 + \frac{n\phi}{\phi_0 + n\phi}\bar{y}.$$

- The prior precision is equal to ϕ_0 and the precision in the likelihood for any y_i is ϕ. Since there are n observations, the precision in the joint likelihood is $n\phi$.
- The posterior mean is a weighted average of the prior mean μ_0 and sample mean $\bar{y}$ where the weights are proportional to the associated precisions.

The posterior accumulates information in the prior and the sample

- The precision of the posterior Normal mean is the sum of the precisions of the prior and likelihood.
- That is,

$$\phi_n = \phi_0 + n\phi.$$

- The implication is that the posterior standard deviation will always be smaller than either the prior standard deviation or the sampling standard error:

$$\sigma_n < \sigma_0, \ \ \sigma_n < \frac{\sigma}{\sqrt{n}}.$$