Comments on Exercises in Chapter 9

9.3

The hard part here is just knowing how much you should say!

For fixed y in bin T_k, we have

$$\begin{aligned}\mathrm{V}(\hat{p}_H(y) &= \mathrm{V}\left(\frac{n_k}{nv_k}\right) \\ &= \frac{1}{(nv_k)^2}\mathrm{V}(n_k).\end{aligned}$$

The probability that any observation falls in the bin T_k is p_k (equation(9.6)). Defining a Bernoulli random variable that takes on a value 1 if an item in the sample falls into T_k and a value of 0, and assuming a random sample, we have n idependent Bernoullis, and their sum, n_k, that is the count of the number that are in T_k. Therefore n_k is binomial with parameters n and p_k, and we have

$$\begin{aligned}\mathrm{V}(\hat{p}_H(y) &= \frac{1}{(nv_k)^2}np_k(1-p_k) \\ &= \frac{p_k(1-p_k)}{nv_k^2}.\end{aligned}$$

9.5

We may be interested in $\mathcal{S}(p')$ because the asymptotic integrated squared bias (and hence the AMISE) of a histogram estimator when the true density is p depends on $\mathcal{S}(p')$. For the gamma, with $\alpha > 0$ and for $0 \leq y$, we have

$$p'(y) = \frac{\alpha-1}{\Gamma(\alpha)\beta^\alpha}y^{\alpha-2}\mathrm{e}^{-y/\beta} - \frac{1}{\Gamma(\alpha)\beta^{\alpha+1}}y^{\alpha-1}\mathrm{e}^{-y/\beta}$$

and

$$\begin{aligned}(p'(y))^2 = &\frac{(\alpha-1)^2}{(\Gamma(\alpha))^2\beta^{2\alpha}}y^{2\alpha-4}\mathrm{e}^{-2y/\beta} \\ &+\frac{1}{(\Gamma(\alpha))^2\beta^{2\alpha+2}}y^{2\alpha-2}\mathrm{e}^{-2y/\beta} \\ &-2\frac{\alpha-1}{(\Gamma(\alpha))^2\beta^{2\alpha+1}}y^{2\alpha-3}\mathrm{e}^{-2y/\beta}.\end{aligned}$$

We now separate the integral into 3 parts, and try to identify a gamma in each. We first of all note that there is a singularity in the gamma integral if $\alpha = 1$, so for the following we require $\alpha \neq 1$.

$$\begin{aligned}\frac{(\alpha-1)^2}{(\Gamma(\alpha))^2\beta^{2\alpha}}\int_0^\infty y^{2\alpha-4}\mathrm{e}^{-2y/\beta}\mathrm{d}y &= \frac{(\alpha-1)^2}{(\Gamma(\alpha))^2\beta^{2\alpha}}\Gamma(2\alpha-3)(\beta/2)^{2\alpha-3}, \\ &= \frac{(\alpha-1)^2\Gamma(2\alpha-3)}{2^{2\alpha-3}(\Gamma(\alpha))^2\beta^3},\end{aligned}$$

$$\begin{aligned}\frac{1}{(\Gamma(\alpha))^2\beta^{2\alpha+2}}\int_0^\infty y^{2\alpha-2}\mathrm{e}^{-2y/\beta}\mathrm{d}y &= \frac{1}{(\Gamma(\alpha))^2\beta^{2\alpha+2}}\Gamma(2\alpha-1)(\beta/2)^{2\alpha-1}, \\ &= \frac{\Gamma(2\alpha-1)}{2^{2\alpha-1}(\Gamma(\alpha))^2\beta^3},\end{aligned}$$

and

$$\begin{aligned}-2\frac{\alpha-1}{(\Gamma(\alpha))^2\beta^{2\alpha+1}}\int_0^\infty y^{2\alpha-3}\mathrm{e}^{-2y/\beta}.\mathrm{d}y &= -2\frac{\alpha-1}{(\Gamma(\alpha))^2\beta^{2\alpha+1}}\Gamma(2\alpha-2)(\beta/2)^{2\alpha-2} \\ &= -\frac{(\alpha-1)\Gamma(2\alpha-2)}{2^{2\alpha-3}(\Gamma(\alpha))^2\beta^3}.\end{aligned}$$

Now, $\mathcal{S}(p')$ is the sum of these three terms, so combine and simplify.

If $\alpha = 1$, $\mathcal{S}(p')$ is not defined.

9.10

First, of all, make sure you can get to the second line of equation (9.33). This requires that $K(u)$ has the properties in equations (9.28) through (9.31), plus a scaling in $K(\cdot)$ so that $\sigma_K^2 = 1$. Also note that equation (9.33) comes from a Taylor series expansion about y; that is, y is a constant, and so is $p(y)$. Keep in mind the structure of the various objects: u, y, and $\nabla p(y)$ are d-vectors; $K(u)$ and $p(y)$ are scalars; V and $\mathrm{H}_p(y)$ are $d \times d$ matrices; and (therefore) $(Vu)^{\mathrm{T}}\mathrm{H}_p(y)Vu$ is a scalar, and $(Vu)^{\mathrm{T}}\mathrm{H}_p(y)VuK(u) = \mathrm{trace}((Vu)^{\mathrm{T}}\mathrm{H}_p(y)VuK(u)$. Another thing to remember is that the trace operation distributes over integration (or summation).

$$\int \mathrm{trace}(f(x))\,\mathrm{d}x = \mathrm{trace}\left(\int f(x)\,\mathrm{d}x\right),$$

where the integration is over a domain in $\mathbb{R}^d$.

$$\begin{aligned}\mathrm{E}\left(\widehat{p}_K(y)\right) &\approx \int_{\mathbb{R}^d} K(u)\left(p(y) - (Vu)^{\mathrm{T}}\nabla p(y) + \frac{1}{2}(Vu)^{\mathrm{T}}\mathrm{H}_p(y)Vu\right)\mathrm{d}u \\ &= p(y)\int_{\mathbb{R}^d} K(u)\,\mathrm{d}u \\ &\quad -(\nabla p(y))^{\mathrm{T}}V\int_{\mathbb{R}^d} uK(u)\,\mathrm{d}u \\ &\quad +\frac{1}{2}\int_{\mathbb{R}^d}\mathrm{trace}\left(u^{\mathrm{T}}V^{\mathrm{T}}\mathrm{H}_p(y)VuK(u)\right)\mathrm{d}u \\ &= p(y) - 0 + \frac{1}{2}\int_{\mathbb{R}^d}\mathrm{trace}\left(V^{\mathrm{T}}\mathrm{H}_p(y)Vuu^{\mathrm{T}}K(u)\,\mathrm{d}u\right) \\ &= p(y) + \frac{1}{2}\mathrm{trace}\left(V^{\mathrm{T}}\mathrm{H}_p(y)V\int_{\mathbb{R}^d} uu^{\mathrm{T}}K(u)\,\mathrm{d}u\right) \\ &= p(y) + \frac{1}{2}\mathrm{trace}\left(V^{\mathrm{T}}\mathrm{H}_p(y)V\sigma_K^2 I\right)\end{aligned}$$

$$= \quad p(y) + \frac{1}{2}\mathrm{trace}\left(V^{\mathrm{T}}\mathrm{H}_p(y)V\right).$$

The expression in the fourth line on page 220 is exactly the same as this for the univariate case, except the assumption that $\sigma_K^2 = 1$ is not used. (Note that $d = 1$, so the integral in that expression is just over $\mathrm{I\!R}$.) So the bias to order h^2 at y is

$$\frac{1}{2}h^2 p''(y)\sigma_K^2,$$

and the asymptotic integrated squared bias is

$$\frac{1}{4}h^4\sigma_K^4 \int (p''(y))^2 \,\mathrm{d}y$$

or

$$\frac{1}{4}h^4\sigma_K^4\mathcal{R}(p).$$

The only thing to do in simplifying equation (3.36) for the univariate case is to note that $|V| = h$. Then combine.

9.11

a. A very straightforward way of doing this is just to use the simple S-Plus statements:

```
x<-c(-1.8,-1.2,-.9,-.3,-.1,.1,.2,.4,.7,1.0,1.3,1.9)
n<-length(x)
# initialize h
h<-0.5
p0<-sum(dnorm((0-x)/h)/(n*h))
```

The result for $h = .5$ is 0.3488.

The same kernel density estimate with a normal kernel can be computed by the S-Plus statement:

```
density(x,n=1,window="gaussian",width=4*h,from=0)
```

It is not clear what algorithm is used in S-Plus. It obviously accumulates some error.

The simple summing method shown above is not efficient if the estimate is to be computed at several points. A method based on an FFT, as mentioned in the text, is better.

b. Do same way.

c. Do same way on a grid of say 100 points between -2 and 2, and then use an S-Plus function like `smooth`.

d. The first thing to remember is that the Hermites are real-valued functions, so the discussion in Section 9.5 does not need to use complex conjugates. Also, note that equations (9.42) and (9.43) assume the orthogonal functions are normalized, and the Hermites given on page 136 are not normalized.