Y, Z, Ryan Quigley
MATH 267A

Problem Set 7

1.

a) In order to show that $\hat{f}_h(a)$ is a probability density function, we must show two things: (1) $\hat{f}_h(a) \geq 0,\ \forall a$ and (2) $\int\limits_{-\infty}^{\infty} \hat{f}_h(a)da = 1$.

For condition (1), we know that n is the sample size so it is a positive integer. It is given that $K(u) \geq 0,\ \forall u$ and h > 0. This implies that $\sum \frac{1}{h}K\left(\frac{a-x_i}{h}\right) \geq 0,\ \forall a$. Thus, we have shown that $\hat{f}_h(a) \geq 0,\ \forall a$.

For condition (2),

$$\begin{aligned}
\int\limits_{-\infty}^{\infty} \hat{f}_h(a)da &= \int\limits_{-\infty}^{\infty} \frac{1}{n}\sum \frac{1}{h}K\left(\frac{a-x_i}{h}\right) da \\
&= \int\limits_{-\infty}^{\infty} \frac{1}{n}\left[\frac{1}{h}K\left(\frac{a-x_1}{h}\right) + ... + \frac{1}{h}K\left(\frac{a-x_n}{h}\right)\right] da \\
&= \frac{1}{n}\left[\int\limits_{-\infty}^{\infty} \frac{1}{h}K\left(\frac{a-x_1}{h}\right) da + ... + \int\limits_{-\infty}^{\infty} \frac{1}{h}K\left(\frac{a-x_n}{h}\right) da\right]
\end{aligned}$$

Let $u_i = \frac{a-x_i}{h} \implies du_i = \frac{1}{h}da$. Note, the transformation does not affect the limits of the integral. By substitution,

$$= \frac{1}{n}\left[\int\limits_{-\infty}^{\infty} K(u_1)du_1 + ... + \int\limits_{-\infty}^{\infty} K(u_n)du_n\right]$$

From the given information, each of the n integrals evaluates to 1.

$$\begin{aligned}
&= \frac{1}{n}(n) \\
&= 1
\end{aligned}$$

Both conditions are met. Therefore, $\hat{f}_h(a)$ is a probability density function. Since our kernel density estimate is a probability density function itself, it can be used to calculate probabilities and make inferences about the underlying population.

b)

$$\begin{aligned}
\int\limits_{-\infty}^{\infty} u\hat{f}_h(u)du &= \int\limits_{-\infty}^{\infty} u \cdot \frac{1}{n}\sum \frac{1}{h}K\left(\frac{u-x_i}{h}\right) du \\
&= \frac{1}{n}\left[\int\limits_{-\infty}^{\infty} u \cdot \frac{1}{h}K\left(\frac{u-x_1}{h}\right) du + ... + \int\limits_{-\infty}^{\infty} u \cdot \frac{1}{h}K\left(\frac{u-x_n}{h}\right) du\right]
\end{aligned}$$

Let $v_i = \frac{u-x_i}{h} \implies dv_i = \frac{1}{h}du \implies u = v_i h + x_i$. Note, the transformation does not affect the limits of the integral. By substitution,

$$
\begin{aligned}
&= \frac{1}{n}\left[\int_{-\infty}^{\infty} (v_1 h + x_1)K(v_1)dv_1 + ... + \int_{-\infty}^{\infty} (v_n h + x_n)K(v_n)dv_n\right] \\
&= \frac{1}{n}\left[h\int_{-\infty}^{\infty} v_1 K(v_1)dv_1 + x_1\int_{-\infty}^{\infty} K(v_1)dv_1 + ... + h\int_{-\infty}^{\infty} v_n K(v_n)dv_n + x_n\int_{-\infty}^{\infty} K(v_n)dv_n\right]
\end{aligned}
$$

From the given information, each of the integrals of the form $\int_{-\infty}^{\infty} K(u)$ evaluates to 1. Additionally, since the kernel is symmetric around 0, the integrals of the form $\int_{-\infty}^{\infty} uK(u)du$ evaluate to 0.

$$
\begin{aligned}
&= \frac{1}{n}[0 + x_1 \cdot 1 + ... + 0 + x_n \cdot 1] \\
&= \frac{x_1 + ... + x_n}{n} \\
&= \bar{x}
\end{aligned}
$$

!!!!INTERPRETATION!!!!

2. The kernel density estimate of $f(2.25)$ is,

$$
\begin{aligned}
\hat{f}_0.8(2.25) &= \frac{1}{5} \cdot \frac{1}{0.8}\left[0 + 0 + 0.09 + 0.39 + 0.09\right] \\
&= 0.14
\end{aligned}
$$

The details of the calculations are summarized in the following table,

i	x_i	$\frac{2.25-x_i}{0.8}$	$\mathbb{1}_{(0,1)}\left(\left\lvert\frac{2.25-x_i}{0.8}\right\rvert\right)$	$\frac{3}{4}\left(1-\left(\frac{2.25-x_i}{0.8}\right)^2\right)$
1	1	1.56	0	-1.08
2	1.2	1.31	0	-0.54
3	1.5	0.94	1	0.09
4	2.8	-0.69	1	0.39
5	3	-0.94	1	0.09