Exercise 6

1. Let $Z_1, \ldots, Z_n$ be independent random variables with distribution $Z_i \sim \mathcal{P}(\lambda_i)$, and let $X = Z_1 + \cdots + Z_n$. Prove that

$$\mathcal{L}(Z_i|X) = \text{Bin}\left(X, \frac{\lambda_i}{\sum_{j=1}^n \lambda_j}\right).$$

2. Let $(x_1, y_1), \ldots, (x_n, y_n)$ be given pairs of real numbers satisfying $0 < x_1 < \ldots < x_n < 1$. We wish to construct a twice continuously differentiable function $f(x)$ on $[0, 1]$ that satisfies

$$f(x_i) = y_i \quad \text{for } i = 1, \ldots, n,$$

and so that the functions is a cubic polynomial on the intervals $[x_i, x_{i+1}]$ for $i = 1, \ldots, n-1$ and a quadratic on the intervals $[0, x_1]$ and $[x_n, 1]$. Show that the number of equations determined by the given conditions is equal to the number of parameters required to specify a function of the stated form.

3. Show that if $K(x)$ is a density function then the kernel density estimator based on $X_1, \ldots, X_n$ given by

$$\widehat{f}_n(x) = \frac{1}{nh} \sum_{i=1}^n f\left(\frac{x - X_i}{h}\right),$$

for some positive bandwidth h, is a density.

4. Using a Taylor expansion, find the order of the bias (in h) of the kernel estimator $\widehat{f}_n(x)$ if the function f has p derivatives.

4. Use orthogonality to explain the decomposition

$$\text{MSE}(\widehat{\theta}) = \text{Bias}^2(\widehat{\theta}) + \text{Var}(\widehat{\theta}).$$

5. When we derived the rate of the bias of the kernel estimator we assumed that the true density f has two derivatives, and that the kernel K is a symmetric density. Under these conditions $\int uK(u)du = 0$, and we obtained that the bias decays at rate h^2.

If you assume that the true density f has higher derivatives, and that $\int u^2K(u)du = 0$, you will get a better rates. Explain and compute this improved rate. Note: if K is symmetric then $\int u^3K(u)du = 0$.