Some of my white board writing from week 3

I discussed in details the main statements (Theorems 2.3, 2.4 and 2.5) from lecture 2.

Theorem 2.3 is presented in sufficient details in the notes and I am not reproducing it again here.

Theorem 2.4 was based on the following two simple observations:

a) If Y is a random variable with $E(Y^2)<\infty$ then the constant a^* that minimizes $E(Y-a)^2 \xrightarrow[a\in R']{}$ is $a^*=E(Y)$.

This is because $\frac{\partial}{\partial a}[E(Y-a)^2]=\frac{\partial}{\partial a}[E(Y^2)-2E(Y)a+a^2]=$
$=-2E(Y)+2a=0$ implies $a^*=E(Y)$ for the stationary point and obviously a^* gives rise to a minimum.

b) If $E|Y|<\infty$ then the constant b^* that minimizes $E|Y-b| \xrightarrow[b\in R']{}$ is $b^*=\text{median}(Y)$.

This is because

$$\frac{\partial}{\partial b}(E|Y-b|)=\frac{\partial}{\partial b}\left[\int_{-\infty}^{b}(b-y)f(y)dy+\int_{b}^{\infty}(y-b)f(y)dy\right]=$$

$$=\frac{\partial}{\partial b}\left[bF(b)-\int_{-\infty}^{b}yf(y)dy+\int_{b}^{\infty}yf(y)dy-b(1-F(b))\right]=$$

$$=F(b)+\cancel{bf(b)}-\cancel{bf(b)}-\cancel{bf(b)}-(1-F(b))+\cancel{bf(b)}=$$

$$=2F(b)-1=0$$ implies $F(b^*)=\frac{1}{2}$ for the stationary point, i.e. b^* is the median. And obviously b^* gives rise to a minimum.

The proof of Theorem 2.5 is also presented in details in the notes.

Then I was illustrating applications of Theorems 2.4 (estimation) and 2.5 (hypothesis testing) in Bayesian context.

Example 2.5.11 about Bayesian estimator of the parameter θ of the Bernoulli distribution when using a Beta(α,β) prior, leading to $\hat{\theta}_{Bayes} = \frac{\sum_{i=1}^{n} X_i + \alpha}{\alpha+\beta+n}$ by following step-by-step the theory from Theorem 2.4 is thoroughly presented in the notes. However, I stressed on the fact that an easier derivation that does not need the calculation of the marginal distribution $g(X)$ of the data, is recommended and is often applied in Bayesian inference.

Namely: knowing that $g(X)$ serves just as a norming constant for the conditional density of $\theta | X$:

$$h(\theta / X) = \frac{f(X|\theta)\tau(\theta)}{g(X)} \propto f(X|\theta)\tau(\theta),$$

we only needed to examine the expression on the top:

$$f(X|\theta)\tau(\theta) = \theta^{\sum_{i=1}^{n} X_i + \alpha - 1}(1-\theta)^{n-\sum_{i=1}^{n} X_i + \beta - 1}$$

This already identifies $h(\theta / X)$ as Beta$\left(\sum_{i=1}^{n} X_i + \alpha, n - \sum_{i=1}^{n} X_i + \beta\right)$

But for any beta distributed random variable Y with parameters α, β, is known that $EY = \frac{\alpha}{\alpha+\beta}$ holds!

Hence we get immediately that $\hat{\theta}_{Bayes} = E(\theta|X) = \frac{\sum_{i=1}^{n} X_i + \alpha}{\cancel{\sum_{i=1}^{n} X_i} + \alpha + n - \cancel{\sum_{i=1}^{n} X_i} + \beta}$

$$= \frac{\sum_{i=1}^{n} X_i + \alpha}{\alpha+\beta+n}.$$

I also notied thas as $n \to \infty$ $\hat{\theta}_{Bayes} = \frac{\bar{X} + \frac{\alpha}{n}}{1 + \frac{\alpha+\beta}{n}}$ becomes very close to $\bar{X}$ as expected.

I also discussed the "approach based on $\propto$" once again in Problem 3 from the Set 1 of tutorial exercises. There we have

$$h(\theta|X) = \frac{f(X|\theta)\,\tau(\theta)}{g(X)} \propto \theta^n e^{-\left(\sum_{i=1}^n x_i + K\right)\theta}.$$

Comparing this conditional density with the Gamma density (the latter being defined as $Y \sim \text{Gamma}(\alpha, \beta)$ having density

$$f(y) = \frac{y^{\alpha-1} e^{-y/\beta}}{\Gamma(\alpha)\,\beta^{\alpha}}, \quad y > 0$$

and $EY = \alpha\beta$) we see immediately that $h(\theta|X)$ must be $\text{Gamma}\left(n+1, \frac{1}{\sum_{i=1}^n x_i + K}\right)$ and hence

$$\hat{\theta}_{Bayes} = \frac{n+1}{\sum_{i=1}^n x_i + K}$$

must hold.

I then discussed Question 6 from the Set 1 to illustrate that the "approach based on $\propto$" also helps a lot in Bayesian hypothesis testing. I discussed in detail on the white board the solution to Question 6. However, the solution is presented very thorougly also in the file containing the solutions to the Set 1 of tutorial exercises. This set is available on moodle hence I abstain from reproducing the solution to Question 6 once again here.