Problem 1

When the context is clear, we write $p(s,t)$ instead of $p_{\epsilon,\nu}(s,t)$.

For $\theta \in \Theta, x \in \mathcal{X}$:

$$\begin{aligned}
p(\theta, x) &= p(\theta, x, T(x)) \\
&= p(x|T(x), \theta)p(\theta, T(x)) \\
&= p(x|T(x))p(\theta, T(x)) && \text{(because } T \text{ is sufficient)} \\
&= \frac{p(x, T(x))}{p(T(x))}p(\theta, T(x)) \\
&= \frac{p(x)}{p(T(x))}p(\theta, T(x)) \\
&= p(x)p(\theta|T(x))
\end{aligned}$$

So that:

$$p(\theta|x) = \frac{p(\theta, x)}{p(x)} = p(\theta|T(x))$$

As a result:

$$H(\theta|X) = \mathbb{E}\log\frac{1}{p(\theta|X)} = \mathbb{E}\log\frac{1}{p(\theta|T)} = H(\theta|T)$$

ie:

$$I(\theta, X) = H(\theta) - H(\theta|X) = H(\theta) - H(\theta|T(X)) = I(\theta, T(X))$$

Problem 2

1.
$$\mathbb{E}\tilde{\alpha}_k = \mathbb{E}[Yf_k(X)] = \mathbb{E}[E[Y|X]f_k(X)] =< f, f_k >= \alpha_k$$
$$\hat{f}_N(x) = \sum_{k\leq N}\mathbb{E}[\tilde{\alpha}_k]f_k(x) = \sum_{k\leq N}\alpha_k f_k(x)$$

$$\mathbb{E}_X B^2(X) = \mathbb{E}|\sum_{k\geq N+1}\alpha_k f_k(X)|^2 = ||\sum_{k\geq N+1}\alpha_k f_k||_X^2 = \sum_{k\geq N+1}\alpha_k^2 \leq \Phi(N)$$

2. (a)
$$E[A(x)] = \sum_{k\leq N}\mathbb{E}Yf_k(X)f_k(x) = \sum_{k\leq N}\alpha_k f_k(x)$$

By Cauchy-Shwarz, $\mathbb{E}[|Yf_k(X)-\alpha_k||Yf_l(X)-\alpha_l|] \leq \sigma_k\sigma_l < \infty$, so $\beta_{k,l} := Cov(Yf_k(X), Yf_l(X)) = \mathbb{E}[(Yf_k(X)-\alpha_k)(Yf_l(X)-\alpha_l)]$ is finite, and smaller than σ^2 in absolute value.

$$\begin{aligned}
\mathrm{var}(A(x)) &= E[|\sum_{k\leq N}(Yf_k(X)f_k(x) - \alpha_k f_k(x))|^2] \\
&= \sum_{k\leq N}\mathbb{E}(Yf_k(X)f_k(x) - \alpha_k f_k(x))^2 \\
&+ \sum_{k,l\leq N, k\neq l}\mathbb{E}(Yf_k(X)f_k(x) - \alpha_k f_k(x))(Yf_l(X)f_l(x) - \alpha_k f_l(x)) \\
&= \sum_{k\leq N}\sigma_k^2 f_k^2(x) + \sum_{k,l\leq N, k\neq l}\beta_{k,l}f_k(x)f_k(x) \\
&\left(\leq \sigma^2(\sum_{k\leq N} f_k(x))^2\right)
\end{aligned}$$

(b)

$$\tilde{f}_N(x) = \frac{1}{n}\sum_{i\leq n}\sum_{k\leq N} Y_i f_k(X_i) f_k(x)$$

$$\begin{aligned}
V(x) &= \mathrm{var}(\tilde{f}_N(x)) \\
&= \frac{1}{n}\,\mathrm{var}\left(\sum_{k\leq N} Yf_k(X)f_k(x)\right) \quad \text{(By independence of the } (X_i, Y_i)\text{)} \\
&= \frac{1}{n}\,\mathrm{var}(A(x)) \\
&\leq \frac{\sigma^2}{n}(\sum_{k\leq N} f_k(x))^2
\end{aligned}$$

So that $\mathbb{E}_X V(X) \leq \frac{\sigma^2}{n}||\sum_{k\leq N} f_k||^2 \leq \frac{N}{n}\sigma^2$

3.

$$\mathbb{E}||\hat{f}_N - f||^2 = \mathbb{E}||\hat{f}_N - \tilde{f}_N||^2 + \mathbb{E}||\tilde{f}_N - f||^2 \leq \frac{N}{n}\sigma^2 + \Phi(N)$$

(a) Setting N to n_0 leads to $\Phi(N) = 0$ and therefore $\mathbb{E}||\hat{f}_N - f||^2 \leq \frac{n_0\sigma^2}{n} = O(\frac{1}{n})$

(b) Setting N to $(\frac{\alpha\lambda n}{\sigma^2})^{\frac{1}{\alpha+1}}$ leads to $\frac{N}{n}\sigma^2+\Phi(N) = \lambda^{\frac{1}{1+\alpha}} n^{\frac{-\alpha}{1+\alpha}} \left(\sigma^2(\frac{\alpha}{\sigma^2})^{\frac{1}{1+\alpha}} + (\frac{\alpha}{\sigma^2})^{\frac{-\alpha}{1+\alpha}}\right) = O(\lambda^{\frac{1}{1+\alpha}} n^{-\frac{\alpha}{1+\alpha}})$

Problem 2

1. Let $r \leq r_0$, then

$$\begin{aligned}
|f_r(x) - f(x)| &= |\int_{B(x,r)} f(y)\frac{dP}{P(B(x,r))} - \int_{B(x,r)} f(x)\frac{dP}{P(B(x,r))}| \\
&\leq \int_{B(x,r)} |f(y) - f(x)|\frac{dP}{P(B(x,r))} \\
&\leq \int_{B(x,r)} ||y - x||\frac{dP}{P(B(x,r))} \\
&\leq r
\end{aligned}$$

2. By the last question, when $r < r_0$: $f_r(x) \leq f(x) + r$. In addition, if $r < f(x)$, then $f_r(x) \leq 2f(x)$

   Furthermore:
$$Q_n(B(x,r)) = \frac{1}{n}\sum_i 1_{|X_i - x| \leq r}$$
   By independence, since $Q(B(x,r)) = E[1_{|X_1-x|\leq r}] = E[1^2_{|X_1-x|\leq r}]$:
$$\operatorname{var} Q_n(B(x,r)) = \frac{1}{n}[Q(B(x,r)) - Q(B(x,r))^2]$$

$$
\begin{aligned}
\operatorname{var} \hat{f}_r(x) &= \frac{1}{P(B(x,r))^2} \operatorname{var} Q_n(B(x,r)) \\
&= \frac{1}{nP(B(x,r))^2}[Q(B(x,r)) - Q(B(x,r))^2] \\
&= \frac{1}{n}[\frac{f_r(x)}{P(B(x,r))} - f_r(x)] \\
&= \frac{1}{n}[\frac{1}{P(B(x,r))} - 1)]f_r(x) \\
&\leq \frac{1}{n}\frac{2f_r(x)}{P(B(x,r))} && \text{(when } r \leq \min(r_0, f(x))\text{)} \\
&\leq \frac{1}{n}\frac{2f_r(x)}{c_0 r^d} && \text{(when } r \leq r_1\text{)} \\
&\leq C\frac{f(x)}{nr^d} && (C = \frac{2}{c_0})
\end{aligned}
$$

The last inequality is valid for all $r \leq \min(f(x), r_0, r_1) := \alpha$.

3. $E\hat{f}_r(x) = \frac{E[Q_n(B(x,r))]}{P(B(x,r))} = f_r(x)$, so

$$
\begin{aligned}
E_{X_Q}[|\hat{f}_r(x) - f(x)|^2] &= E_{X_Q}[|\hat{f}_r(x) - f_r(x)|^2] + |f_r(x) - f(x)|^2 \\
&\leq \frac{2f(x)}{c_0 n r^d} + r^2 && \text{(for } r < \alpha\text{)} \\
&\leq \frac{a}{r^d} + r^2 && (a := \frac{2f(x)}{c_0 n})
\end{aligned}
$$

Let's define $g : r \to \frac{a}{r^d} + r^2$, $g'(r) = \frac{-ad}{r^{d+1}} + 2r$, $g''(r) = \frac{ad(d+1)}{r^{d+2}} + 2 > 0$,

g is convexe, so it has a global minimum when $g'(r) = 0$, ie $r = \left(\frac{ad}{2}\right)^{\frac{1}{d+2}}$ and in that case

$$
\begin{aligned}
g(r) &= a^{1-\frac{d}{d+2}}(\frac{d}{2})^{\frac{-d}{d+1}} + (\frac{ad}{2})^{\frac{2}{d+2}} \\
&= a^{\frac{2}{d+2}}\left((\frac{d}{2})^{\frac{-d}{d+1}} + (\frac{d}{2})^{\frac{2}{d+2}}\right) \\
&= n^{-\frac{2}{d+2}} \underbrace{(\frac{2f(x)}{c_0})^{\frac{2}{d+2}}\left((\frac{d}{2})^{\frac{-d}{d+1}} + (\frac{d}{2})^{\frac{2}{d+2}}\right)}_{C}
\end{aligned}
$$

We define $r(n) = \left(\frac{f(x)d}{nc_0}\right)^{\frac{1}{d+2}}$. This quantity is smaller than α, when $n > \frac{f(x)d}{c_0\alpha^{d+2}}$, and then we have that:

$$E_{X_Q}[|\hat{f}_r(x) - f(x)|^2] \leq Cn^{\frac{-2}{d+2}}$$

Problem 4

1. The X_i are iid and $E[\frac{1_{|X_i-x|\leq r}}{P(B(x,r))}] = \frac{Q(B(x,r))}{P(B(x,r))}$, var $\frac{1_{|X_i-x|\leq r}}{P(B(x,r))} = \frac{Q(B(x,r))-Q^2(B(x,r))}{P(B(x,r))}$ CLT:

$$\sqrt{n}(\hat{f}_r(x)-f_r(x)) = \frac{1}{\sqrt{n}}\sum_{i\leq n}\frac{1_{|X_i-x|\leq r}}{P(B(x,r))}-\frac{Q(B(x,r))}{P(B(x,r))} \xrightarrow{d} \mathcal{N}(0, \frac{Q(B(x,r)) - Q^2(B(x,r))}{P(B(x,r))})$$

Since $\frac{Q_n(B(x,r))-Q_n^2(B(x,r))}{P(B(x,r))} \xrightarrow{\text{a.s.}} \frac{Q(B(x,r))-Q^2(B(x,r))}{P(B(x,r))}$, by slutsky:

$$\sqrt{\frac{nP(B(x,r))}{Q_n(B(x,r)) - Q_n^2(B(x,r))}}(\hat{f}_r(x) - f_r(x)) \xrightarrow{d} \mathcal{N}(0,1)$$

If z_α the $1-\frac{\alpha}{2}$ quantile of the normal distribution, then:

$$\mathbb{P}\left(\sqrt{\frac{nP(B(x,r))}{Q_n(B(x,r)) - Q_n^2(B(x,r))}}|\hat{f}_r(x) - f_r(x)| \leq z_\alpha\right) \to 1-\alpha$$

Therefore $[\hat{f}_r(x) \pm z_\alpha\sqrt{\frac{Q_n(B(x,r))-Q_n^2(B(x,r))}{nP(B(x,r))}}]$ is a $1-\alpha$ confidence interval for $f_r(x)$.

2. Under the new assumption, $f_r(x) - r \leq f(x) \leq r + f_r(x)$, so

$$[\hat{f}_r(x) \pm \left(r + z_\alpha\sqrt{\frac{Q_n(B(x,r)) - Q_n^2(B(x,r))}{nP(B(x,r))}}\right)]$$

is a $1-\alpha$ confidence interval for $f(x)$.

Problem 5

Lemma 0.1. *Let* $X_n \to \mathcal{N}(0, \sigma_x^2)$*,* $Y_n \to \mathcal{N}(0, \sigma_y^2)$*,* X_n *and* Y_n *are independent for all* n*. Then* $(X_n, Y_n) \to \mathcal{N}(0, diag(\sigma_x^2, \sigma_y^2))$ *And as a corrolary, using the continuous mapping theorem:* $X_n + Y_n = (1,1)'(X_n, Y_n) \to \mathcal{N}(0, \sigma_x^2 + \sigma_y^2)$

Proof. Let F_n (resp. G_n) be the c.d.f of X_n (resp Y_n), and let H_n be the joint c.d.f of (X_n, Y_n). Let F (resp. G) be the cdf of $N(0, \sigma_x^2)$ (resp $N(0, \sigma_y^2)$), and H be the c.d.f of $N(0, diag(\sigma_x^2, \sigma_y^2))$.

By independence: $H_n(x, y) = F_n(x)G_n(y) \to_n F(x)G(y) = H(x, y)$. Which proves the lemma. □

Let's call $p := R(h), q := R(h')$. H_0 is equivalent to $p = q$, and H_1 is equivelent to $p \neq q$.

- We know that $1_{X=h(Y)}$ is binomial of parameter p. Same for h'.
- $E[\hat{R}(h) + \hat{R}'(h')] = p + 1$. By independence of all the X_i, X_i',
- Law of large numbers: $\hat{R}(h) \xrightarrow{as} p$, $\hat{R}(h') \xrightarrow{as} q$
- CLT:

$$\sqrt{n}(\hat{R}(h) - p) \xrightarrow{d} \mathcal{N}(0, p(1-p))$$
$$\sqrt{n}(\hat{R}'(h) - q) \xrightarrow{d} \mathcal{N}(0, q(1-q))$$

  By independence of X_i and X_i' for all i, using the lemma:

$$\sqrt{n}(\hat{R}(h') - \hat{R}'(h') - (p-q)) \xrightarrow{d} \mathcal{N}(0, p(1-p) + q(1-q))$$

- By Slutsky:

$$\sqrt{\frac{n}{\hat{R}(h)(1-\hat{R}(h)) + \hat{R}(h')(1-\hat{R}(h')}}(\hat{R}(h') - \hat{R}'(h') - (p-q)) \xrightarrow{d} \mathcal{N}(0,1)$$

- $P(\sqrt{\frac{n}{\hat{R}(h)(1-\hat{R}(h)) + \hat{R}(h')(1-\hat{R}(h')}}|\hat{R}(h') - \hat{R}'(h') - (p-q)| \leq z_\alpha) \to 1 - \alpha$

So under H_0:

$$P(\sqrt{\frac{n}{\hat{R}(h)(1-\hat{R}(h))+\hat{R}(h')(1-\hat{R}(h')}}|\hat{R}(h')-\hat{R}'(h')|\leq z_\alpha)\to 1-\alpha$$

The test whose rejection region is the following

$$1\left\{\left(\frac{n}{\hat{R}(h)(1-\hat{R}(h))+\hat{R}(h')(1-\hat{R}(h')}\right)^{\frac{1}{2}}|\hat{R}(h')-\hat{R}'(h')|>z_\alpha\right\}$$

is of level $1-\alpha$.

Problem 6

Let f_0 (resp f_1) be the density of Z_0 (resp Z_1)

$$\begin{aligned}
P(h(X)\neq Y) &= E[1_{h(X)\neq Y}]\\
&= E[1_{h(X)\neq Y}|Y=0]P(Y=0)+E[1_{h(X)\neq Y}|Y=1]P(Y=1)\\
&= \frac{1}{2}(P(h(Z_0)\neq 0)+P(h(Z_1\neq 1))\\
&= \frac{1}{2}(\int_R f_0(z)1_{h(z)\neq 0}+\int_R f_1(z)1_{h(z)\neq 1})\\
&\geq \frac{1}{2}(\int_R \min(f_0(z),f_1(z))) && \text{(since } h(z)\in\{0,1\})\\
&= \frac{1}{2}(\int_R f_0(z)1_{f_0(z)<f_1(z)}+f_1(z)1_{f_0(z)\geq f_1(z)})\\
&= \frac{1}{2}(\int_R f_0(z)1_{g(z)\neq 0}+f_1(z)1_{g(z)\neq 1}) && (g(z)=1_{f_0(z)<f_1(z)})\\
&= R(g)=\inf_h R(h)\\
&= \frac{1}{2}(P(Z_0\in R_0)+P(Z_1\in R_1))
\end{aligned}$$

Where

$$R_0=g^{-1}(1)=\{z|\lambda_0e^{-\lambda_0 z}\leq\lambda_1e^{-\lambda_1 z}\}=(-\infty,\frac{1}{\lambda_1-\lambda_0}\log\frac{\lambda_1}{\lambda_0})$$

$$R_1=[\frac{1}{\lambda_1-\lambda_0}\log\frac{\lambda_1}{\lambda_0},\infty)$$

Problem 7

Let $\hat{f}$ be fixed.

While for each $P_{X,Y}$, the quantity $\mathbb{E}\bar{R}(h_{\hat{f}})/a_n$ converges to 0, the convergence can occur at different speeds so that at any moment n, there might exist a $P_{X,Y}$ for which $\mathbb{E}\bar{R}(h_{\hat{f}})/a_n$ is still larger than c.

If such situtation happens, $\sup_{P_{X,Y}} \mathbb{E}\bar{R}(h_{\hat{f}})/a_n \geq c$, and taking the infimum over all $\hat{f}$ and taking the $\liminf$, we get that a_n is the minimax rate for P.