q4

Suppose $\tau_n \uparrow \infty$ and for all $\epsilon > 0$, there exists $B > 0$, such that

$$\inf_n Pr(|\tau_n(\hat{\theta} - \theta)| \leq B) \geq 1 - \epsilon$$

Equivalently, we have $\inf_n Pr(|\hat{\theta} - \theta| \leq \frac{B}{|\tau_n|}) \geq 1 - \epsilon$.
Now, we can choose some N $\in \mathbb{N}$ such that, for all $n > N$, $\frac{B}{\tau_n} < \delta$ as B is a constant and $\tau_n \uparrow \infty$. Then, we have that, for all $n > N$,

$$\begin{aligned}
1 - \epsilon &\leq \inf_n (Pr(|\hat{\theta} - \theta)| \leq \frac{M}{|\tau_n|}) \\
&\leq \inf_{n>N} (Pr(|\hat{\theta} - \theta)| \leq \frac{M}{|\tau_n|}) \\
&\leq \inf_{n>N} (Pr(|\hat{\theta} - \theta)| \leq \delta)
\end{aligned}$$

This equivalently states that tightness of $\tau_n(\hat{\theta} - \theta)$ implies that
$Pr(|\hat{\theta} - \theta)| \leq \delta) \to 1$

q8

0.1 a

Noting that $f(y|x) = 0$ if $f_X = 0$, we know the integral over $\mathbb{R}^{\Bbbk} \times \mathbb{R}$ simplifies to the integral over the area where $f_X(x) > 0$ (as it is 0 everywhere else).

$$\begin{aligned}
E[m^{*2}(X)] &= \int (\int y f(y|x) dy)^2 f_X(x) dx \\
&\leq \int (\int |y| \frac{f(y,x)}{f_X(x)} dy)^2 f_X(x) dx
\end{aligned}$$

Knowing that $\int \frac{f(y,x)}{f_X(x)} dy = 1$, we know (i.e by Cauchy -Schwartz):

$$\int y \frac{f(y,x)}{f_X(x)} dy \leq (\int y^2)^{.5} (\int \frac{f(y,x)}{f_X(x)} dy)^{.5}$$

Thus, we can write out

$$\begin{aligned}
E[m^{*2}(X)] &\leq \int (\int y^2 \frac{f(y,x)}{f_X(x)} dy) f_X(x) dx \\
&= \int \int (y^2 \frac{f(y,x)}{f_X(x)} f_X(x)) dy dx \\
&\leq \int \int y^2 f(x,y) dy dx \leq E(Y^2) < \infty
\end{aligned}$$

as, again, f_X is zero everywhere else.

0.2 b

Recall, from class that

$$\begin{aligned} E[(y-m(x))^2] =& E[(y-m(x)+m^*(x)-m^*(x))^2] \\ =& E[(y-m^*(x))^2]+2E[(y-m^*(x))(m^*(x)-m(x))]+E[(m^*(x)-m(x))^2] \\ \geq& E[(Y-m^*(X))^2] \end{aligned}$$

Thus, we found that $\min E[(Y-m^*(X))] \Leftrightarrow E[(Y-m^*(X))m(X)]=0$ for all m(X). Now, see that

$$\begin{aligned} E[(y-m^*(x))m(x)] &= \int\int(y-m^*(x))m(x)f(y,x)dydx \\ &= \int(\int(y-m^*(x))m(x)f(y,x)dy)dx \\ &= \int m(x)f_X(x)(\int yf(y|x)-m^*(x)f(y|x)dy)dx \\ &= \int m(x)m^*(x)f_X(x)dx - \int m(x)m^*(x)(\int f(y|x)dy)f_X(x)dx \end{aligned}$$

As $\int f(y|x)dy$ just integrates to 1, these two terms on the left and right are equal (namely $E[(y-m^*(x))m(x)]=0$

q12

0.3 a

Take

$$Y=\beta_0+\beta_1 X+U$$

Now, consider

$$\begin{aligned} \beta_1 =& \frac{Cov(X,Y)}{\sigma_X^2} && (1) \\ =& \rho_{X,Y}\frac{\sigma_Y}{\sigma_X} && (2) \end{aligned}$$

Thus, the $|\beta_1|<1$ does not necessarily mean $\frac{Var(X)}{Var(Y)}<1$ as we need $\frac{\beta_1}{\rho_{X,Y}}<1$. Note, you can also see that $|\beta|<1$ doesn't imply the claim from just writing out

$$\frac{var(Y)}{var(X)}=\frac{\beta_1^2 var(X)+var(U)}{var(X)}$$

0.4 b

As $\sigma_X=\sigma_Y$, the above equation (2) implies that we have $\beta_1=\rho_{X,Y}$, so $\beta_1=1$ iff $\rho_{X,Y}=1$. Also, as $\sigma_Y^2=\beta_1^2\sigma_X^2+\sigma_U^2$, we require that, if $Cov(X,U)=0$, $\sigma_U^2=0$

0.5 c

Again, as we have

$$\begin{aligned}\beta_1 &= \rho_{X,Y}\frac{\sigma_Y}{\sigma_X}\\ &= \rho_{X,Y}\frac{\sigma_X}{\sigma_Y}\\ &= \alpha_1\end{aligned}$$

as the distributions (and variances) are equal. The equality of α_1 and β_1 requires, either $\rho_{X,Y} = 0$ or $\sigma_X = \sigma_Y$

q16

Intuitively, we have that since E(V) = 0 and $V \in \{0, 1\}$, we cannot have the measurement error to "cancel out" in the case of classical measurement error as if X =1, the measurement error must be negative and if X = 0, the measurement error must be positive, so it must be negatively correlated with X.
Note that if E(V) = 0,

$$\begin{aligned}Cov(X,V) &= E((X - E(X))(V - E(V)))\\ &= E(XV) - E(X)E(V)\\ &= E(XV)\end{aligned}$$

Now, looking at variance of $\hat{X}$, we see that if $Cov(X,V) = E(XV) = 0$, $Var(\hat{X}) = Var(X)$

$$\begin{aligned}Var(\hat{X}) =& E(X^2) + E(V^2) + E(XV) - E(\hat{X})^2\\ =& E(X^2) - E(X)^2 + E(V^2)\\ =& Var(X) + Var(V)\end{aligned}$$

Here, as $E(X^2) = E(X)$,

$$var(\hat{X}) = E(\hat{X})(1 - E(\hat{X})) = var(X)$$

so $Var(V) = 0$ so V= 0 and $\hat{X}$ is just X.