1 Question 3

In what sense is $o_P(1) = \mathcal{O}_P(1)$? Is $\mathcal{O}_P(1) = o_p(1)$?

We say that a sequence of random variables, X_n, is $o_P(1)$ if $X_n \xrightarrow{p} 0$. We say that X_n is $\mathcal{O}_P(1)$ if X_n is tight. Since we have that

$$X_n = o_P(1) \implies X_n \xrightarrow{d} 0$$

and

$$X_n \xrightarrow{d} X \implies X_n = \mathcal{O}_P(1),$$

(where X is a random variable) we have,

$$X_n = o_P(1) \implies X_n = \mathcal{O}_P(1).$$

In this sense,

$$o_P(1) = \mathcal{O}_P(1).$$

However, the converse is not true in general. For instance, realize that $X_n \xrightarrow{d} X$ is a sufficient condition for tightness, but not for convergence in probability. Only when X is a constant does it imply convergence in probability, but even then, X must equal 0 for $X_n = o_P(1)$.

An even stronger statement can be said though: in general, tightness does not imply convergence in distribution, and therefore does not imply convergence in probability. Consider, a sequence of random variables, X_n, where $X_{2n} \sim U[0,1]$, and $X_{2n+1} \sim U[2,3]$. It is obvious that X_n does not converge in distribution. However, it is tight. To prove this, take $M_\epsilon = 3$. Then, we have

$$\sup \Pr(|X_n| > 3) < \epsilon, \forall \epsilon > 0.$$

Thus, $X_n = \mathcal{O}_P(1)$, but $X_n \neq o_P(1)$.

2 Question 7

To prove this, notice first that it is essentially Jensen's Inequality with conditional expectations. Thus, we will need the Chordal Slope Lemma. Also, (after defining $c := \mathbf{E}[Y|X]$) the following objects will be helpful:

$$\begin{aligned}
\Delta_{+,h(c)} &:= \frac{f(c+h) - f(c)}{h} \\
\Delta_{-,h(c)} &:= \frac{f(c) - f(c-h)}{h} \\
D_+(c) &:= \lim_{h \downarrow 0} \Delta_{+,h(c)} \\
D_-(c) &:= \lim_{h \downarrow 0} \Delta_{-,h(c)},
\end{aligned}$$

where f is a convex function. It is also easy to see by the Chordal Slope Lemma that $D_-(c)$ and $D_+(c)$ are bounded below and above respectively by $\Delta_{-,h(c)}$ and $\Delta_{+,h(c)}$.

Next, select an $m \in [D_-(c), D_+(c)]$, and define

$$L(x) := f(c) + m(x - c).$$

We we now want to show that $L(x) \leq f(x)$. There are three cases: when $c > x$, $c = x$, and when $c < x$. From this point on, we will replace the previous convex function f with another convex function, call it ϕ.

First consider $c = x$. The inequality holds trivially.

Next, take $c > x = c - h$. Notice that since $m \in [D_-(c), D_+(c)]$, we get:

$$\begin{aligned} m &\geq \frac{\phi(c) - \phi(x)}{c - x} \\ \phi(c) + m(x - c) &\leq \phi(x) \\ L(x) &\leq \phi(x). \end{aligned}$$

For the last case, take $c < x = c - h$. Just like above, we get:

$$\begin{aligned} m &\leq \frac{\phi(x) - \phi(c)}{x - c} \\ \phi(c) + m(x - c) &\leq \phi(x) \\ L(x) &\leq \phi(x). \end{aligned}$$

Thus, $L(x) \leq \phi(x)$.

Next, take, $x = Y$ and recall that $c := \mathbf{E}[Y|X]$. We have that

$$\begin{aligned} L(Y) &\leq \phi(Y) \\ 0 &\leq \phi(Y) - L(Y) \\ 0 &\leq \mathbf{E}[\phi(Y) - L(Y)|X] && 3) \\ 0 &\leq \mathbf{E}[\phi(Y)|X] - \mathbf{E}[L(Y)|X] && 1) \\ \mathbf{E}[L(Y)|X] &\leq \mathbf{E}[\phi(Y)|X] \\ \mathbf{E}[\phi(\mathbf{E}[Y|X])|X] + \mathbf{E}[mY|X] - \mathbf{E}[m\mathbf{E}[Y|X]|X] &\leq \mathbf{E}[\phi(Y)|X] && 1) \\ \phi(\mathbf{E}[Y|X]) + m\mathbf{E}[Y|X] - m\mathbf{E}[Y|X] &\leq \mathbf{E}[\phi(Y)|X] && 1)\ \&\ 2) \\ \phi(\mathbf{E}[Y|X]) &\leq \mathbf{E}[\phi(Y)|X]. \end{aligned}$$

And thus, our result has been obtained. The steps above can be justified from two properties of conditional expectation (the steps have been labeled accordingly). Namely: 1) $\mathbf{E}[Y + Z|X] = \mathbf{E}[Y|X] + \mathbf{E}[Z|X]$; 2) If $Y = f(X)$, then $\mathbf{E}[Y|X] = f(X)$; and 3) we know that if $\Pr(0 \leq Y) = 1$, then $\Pr(0 \leq \mathbf{E}[Y|X]) = 1$.

3 Question 11

To answer this question, we are going to need to prove the following fact: that independence of X and Y implies that $\mathbf{E}[Y|X] = \mathbf{E}[Y]$, which is a constant.

Consider the definition of conditional expectation. Since all we are given is that the first moment for Y exists, we have to work from the following definition: $\mathbf{E}[Y|X]$ is any $m^*(X)$ with $\mathbf{E}[|m^*(X)|] < \infty$ such that for any Borel set B in $\mathcal{B} \subset \mathbb{R}^k$,

$$\mathbf{E}[(Y - m^*(X))\mathbb{1}_{\{X \in B\}}] = 0.$$

Working from this definition, we can obtain our result. First, let $m^*(X) = \mathbf{E}[Y]$ and B an arbitrary Borel set, then test to see if it solves the following:

$$\begin{aligned}
\mathbf{E}[(Y - m^*(X))\mathbb{1}_{\{X \in B\}}] &= 0 \\
\mathbf{E}[(Y - \mathbf{E}[Y])\mathbb{1}_{\{X \in B\}}] &= 0 \\
\mathbf{E}[Y\mathbb{1}_{\{X \in B\}}] &= \mathbf{E}[\mathbf{E}[Y]]\mathbf{E}[\mathbb{1}_{\{X \in B\}}] \\
\mathbf{E}[Y]\mathbf{E}[\mathbb{1}_{\{X \in B\}}] &= \mathbf{E}[\mathbf{E}[Y]]\mathbf{E}[\mathbb{1}_{\{X \in B\}}] && \text{by } Y \perp\!\!\!\perp X \\
\mathbf{E}[Y]\mathbf{E}[\mathbb{1}_{\{X \in B\}}] &= \mathbf{E}[Y]\mathbf{E}[\mathbb{1}_{\{X \in B\}}] \\
\mathbf{E}[Y]\Pr\{X \in B\} &= \mathbf{E}[Y]\Pr\{X \in B\}.
\end{aligned}$$

Since $\mathbf{E}[Y]$ works above, and $\mathbf{E}[Y|X] := m^*(X)$ we have that $\mathbf{E}[Y|X] = \mathbf{E}[Y]$. Thus, $\mathbf{E}[Y|X]$ is equal to a constant with probability one, and that constant is $\mathbf{E}[Y]$.

4 Question 15

We are given that $\mathbf{E}[Y|X] = X'\beta$, and that $Y = X'\beta + U$. This implies that $\mathbf{E}[U|X] = 0$. To see this take the conditional expectation of $Y = X'\beta + U$:

$$\begin{aligned}
\mathbf{E}[Y|X] &= \mathbf{E}[X'\beta + U|X] \\
\mathbf{E}[Y|X] &= \mathbf{E}[X'\beta|X] + \mathbf{E}[U|X] && 1) \\
\mathbf{E}[Y|X] &= X'\beta + \mathbf{E}[U|X]. && 2)
\end{aligned}$$

And since we are given that $\mathbf{E}[Y|X] = X'\beta$, it is immediate that:

$$\mathbf{E}[U|X] = 0.$$

As in Question 7, the steps above can be justified from two properties of conditional expectation (the steps have been labeled accordingly). Namely: 1) $\mathbf{E}[Y + Z|X] = \mathbf{E}[Y|X] + \mathbf{E}[Z|X]$; and 2) If $Y = f(X)$, then $\mathbf{E}[Y|X] = f(X)$.

Although this implies that U is mean independent of X, it does not imply independence. Notice that because Y takes values in $\{0, 1\}$, we have that $Y|X$ is Bernoulli with $p = \mathbf{E}[Y|X]$, i.e.

$$Var[Y|X] = \mathbf{E}[Y|X](1 - \mathbf{E}[Y|X]).$$

We can also observe that $Var[U|X] = Var[Y|X] = \mathbf{E}[Y|X](1 - \mathbf{E}[Y|X])$. And since it is given that $\mathbf{E}[Y|X] = X'\beta$, we have that:

$$Var[U|X] = X'\beta(1 - X'\beta)$$

which does depend on X.