Problem 1. [20 points]
The statistic $Z$ can be rewritten as:

$$Z = \frac{\hat{\theta} - \theta_1}{\text{se}} + \frac{\theta_1 - \theta_0}{\text{se}}$$

When the true $\theta = \theta_1 > \theta_0$ we have that for large $n$, $Z \approx N \left( \frac{\theta_1 - \theta_0}{\text{se}}, 1 \right)$, such that the power of the test is given by:

$$\beta(\theta_1) = \mathbb{P}_{\theta_1}(Z < -z_{\alpha/2}) + \mathbb{P}_{\theta_1}(Z > z_{\alpha/2})$$

$$\approx \Phi \left( -z_{\alpha/2} - \frac{\theta_1 - \theta_0}{\text{se}} \right) + 1 - \Phi \left( z_{\alpha/2} - \frac{\theta_1 - \theta_0}{\text{se}} \right)$$

where $\Phi$ is the CDF for the standard normal. Since $\frac{\theta_1 - \theta_0}{\text{se}} = \sqrt{nI(\hat{\theta})} \left( \theta_1 - \theta_0 \right) \to \infty$, then $\Phi \left( a - \frac{\theta_1 - \theta_0}{\text{se}} \right) \to 0$ as $n \to \infty$, for any $a \in \mathbb{R}$. This proves the claim. The case $\theta_1 < \theta_0$ can be treated similarly.

Alternate solution.

Under $\theta = \theta_1$, $\frac{Z}{\sqrt{n}} = \frac{\hat{\theta} - \theta_0}{\sqrt{1/I(\theta)}} \xrightarrow{P} (\theta_1 - \theta_0)\sqrt{I(\theta_1)}$ by using continuous mapping theorem.

Hence for $\forall \epsilon > 0$, $P_{\theta_1} \left( \left| \frac{Z}{\sqrt{n}} - (\theta_1 - \theta_0)\sqrt{I(\theta_1)} \right| < \epsilon \right) \to 1$, and we can choose in particular $\epsilon = \frac{1}{2} |\theta_1 - \theta_0|\sqrt{I(\theta_1)}$.

Then since $P_{\theta_1} \left( |Z| > \frac{1}{2} \sqrt{n}|\theta_1 - \theta_0|\sqrt{I(\theta_1)} \right) \geq P_{\theta_1} \left( \left| \frac{Z}{\sqrt{n}} - (\theta_1 - \theta_0)\sqrt{I(\theta_1)} \right| < \frac{1}{2} |\theta_1 - \theta_0|\sqrt{I(\theta_1)} \right)$,

so $\lim_{n \to \infty} P_{\theta_1} \left( |Z| > \frac{1}{2} \sqrt{n}|\theta_1 - \theta_0|\sqrt{I(\theta_1)} \right) = 1$.

So if we let $N$ be s.t. $\frac{1}{2} \sqrt{N}|\theta_1 - \theta_0|\sqrt{I(\theta_1)} > z_{\frac{\alpha}{2}}$, then $\forall n \geq N$, $P_{\theta_1} \left( |Z| > z_{\frac{\alpha}{2}} \right) \geq P_{\theta_1} \left( |Z| > \frac{1}{2} \sqrt{n}|\theta_1 - \theta_0|\sqrt{I(\theta_1)} \right)$, so $\lim_{n \to \infty} P_{\theta_1} \left( |Z| > z_{\frac{\alpha}{2}} \right) = 1$.

$\therefore \lim_{n \to \infty} \beta(\theta_1) = \lim_{n \to \infty} P_{\theta_1} \left( |Z| > z_{\frac{\alpha}{2}} \right) = 1$.
Problem 2. [30 points]
For the Wald test, if $X \sim \text{Bin}(n, p)$ then the likelihood function is:

$$\mathcal{L}(p) \propto p^X (1 - p)^{n - X}$$

Taking the log:

$$\log \mathcal{L}(p) = C + X \log(p) + (n - X) \log(1 - p)$$

Calculating its first derivative:

$$\frac{\partial}{\partial p} \log \mathcal{L}(p) = \frac{X}{p} + \frac{n - X}{1 - p}$$

Setting it equal to 0 we obtain the MLE $\hat{p} = X/n$. This point is the global maximum since:

$$\frac{\partial^2}{\partial p^2} \log \mathcal{L}(p) = -\left(\frac{X}{p^2} + \frac{n - X}{(1 - p)^2}\right) < 0$$

Now since $\mathbb{V}[\hat{p}] = \frac{\hat{p}(1 - \hat{p})}{n}$, then $\hat{se} = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}$. So the wald statistic is:

$$Z = \frac{\hat{p} - p}{\hat{se}} = \frac{\sqrt{n}(\hat{p} - p)}{\sqrt{\hat{p}(1 - \hat{p})}} \sim N(0, 1)$$

so at level $\alpha$ the rejection region is $R = \{|Z| > z_{\alpha/2}\}$.

The notation in the book is different from the notation in the notes. Here we follow the book and let:

$$\lambda(X) = 2 \log \left(\frac{\sup_{\theta \in \Theta} \mathcal{L}(\theta)}{\sup_{\theta \in \Theta_0} \mathcal{L}(\theta)}\right) = 2 \log \left(\frac{\mathcal{L}(\hat{\theta})}{\mathcal{L}(\hat{\theta}_0)}\right)$$

By the first part of the problem we know that $\hat{\theta} = \hat{p} = X/n$, also $\hat{\theta}_0 = p_0$ so:

$$\lambda(X) = 2n \left(\hat{p} \log \left(\frac{\hat{p}}{p_0}\right) + (1 - \hat{p}) \log \left(\frac{1 - \hat{p}}{1 - p_0}\right)\right)$$

Again by asymptotic theory we know that $\lambda(X) \sim \chi^2_1$. Thus at level $\alpha$ the rejection region is $R = \{\lambda > \chi^2_1, \alpha\}$.

By Continuous Mapping Theorem we have that $Z \sim N(0, 1)$ implies $Z^2 \sim \chi^2_1$ such that the Wald test and the likelihood ratio test are asymptotically equivalent. Moreover, also simply notice that in both tests we reject $H_0$ when $\hat{p}$ is very small or very large.
Problem 3. [30 points]
Again we use the notation from the book, let:

$$\begin{align*}
\lambda &= 2 \log \left( \sup_{\theta \in \Theta} \frac{\mathcal{L}(\theta)}{\mathcal{L}(\hat{\theta})} \right) = 2 \log \left( \frac{\mathcal{L}(\hat{\theta})}{\mathcal{L}(\theta_0)} \right) = 2(\log \mathcal{L}(\hat{\theta}) - \log \mathcal{L}(\theta_0))
\end{align*}$$

since $\Theta_0 = \{\theta_0\}$ implies $\hat{\theta}_0 = \theta_0$. Let $l(\theta) = \log \mathcal{L}(\theta)$. In what follows we assume $H_0$ is true, i.e. $\theta_0$ is the true value, because we want to show $W^2 \xrightarrow{\mathcal{L}} 1$ under the null hypothesis. Following the hint, we expand $l(\theta_0)$ by Taylor around point $\theta_0 = \hat{\theta}$, where $\hat{\theta}$ is the MLE:

$$l(\theta_0) = l(\hat{\theta}) + l'(\hat{\theta})(\theta_0 - \hat{\theta}) + \frac{1}{2} l''(\hat{\theta})(\theta_0 - \hat{\theta})^2 + \frac{1}{6} l'''(\theta^*)(\theta_0 - \hat{\theta})^3,$$

with $\theta^* = \hat{\theta} + \tau(\theta_0 - \hat{\theta})$, $\tau \in (0, 1)$

The “regularity conditions” for consistency and asymptotic normality of MLE are supposed to be satisfied (otherwise we could not deal with $W$) and implies that $l'''(\theta) = O_p(n)$, for any $\theta \in (\theta_0 - c, \theta_0 + c)$, for some constant $c > 0$ (see for example Casella & Berger p. 516). Thus:

$$\begin{align*}
\lambda &= -l''(\hat{\theta}) (\theta_0 - \hat{\theta})^2 + \frac{1}{3} \underbrace{l'''(\theta^*)}_{O_p(n)} (\theta_0 - \hat{\theta})^3 \\
&= (\theta_0 - \hat{\theta})^2 \left(-l''(\hat{\theta}) + o_p(n)\right) \\
&= \left(\sqrt{n}(\theta_0 - \hat{\theta})\right)^2 \left(-1 \frac{1}{n} l''(\hat{\theta}) + o_p(1)\right)
\end{align*}$$

The Wald statistic is $W = \sqrt{n}(\theta_0 - \hat{\theta}) \sqrt{\mathcal{I}(\hat{\theta})}$. Thus we get:

$$\begin{align*}
\frac{W^2}{\lambda} &= \left(\sqrt{n}(\theta_0 - \hat{\theta})\right)^2 \frac{\mathcal{I}(\hat{\theta})}{\left(-\frac{1}{n} l''(\hat{\theta}) + o_p(1)\right) \mathcal{I}(\theta_0)} = \left(-\frac{1}{n} l''(\hat{\theta}) + o_p(1)\right) \\
&\xrightarrow{\mathcal{L}} \frac{\mathcal{I}(\hat{\theta})}{\mathcal{I}(\theta_0)} \\
&\xrightarrow{\mathcal{L}} 1
\end{align*}$$

By equivariance, $\mathcal{I}(\hat{\theta})$ is the MLE of $\mathcal{I}(\theta_0)$, then $\mathcal{I}(\hat{\theta}) \xrightarrow{\mathcal{L}} \mathcal{I}(\theta_0)$. As seen in the proof of Theorem 8 of Lecture Notes 10, we have $-\frac{1}{n\mathcal{I}(\theta_0)} l''(\hat{\theta}) \xrightarrow{\mathcal{L}} 1$ by WLLN (see details below). Thus, by Slutsky’s Theorem and Continuous Mapping Theorem we get:

$$\begin{align*}
\frac{W^2}{\lambda} &= \frac{\mathcal{I}(\hat{\theta})}{\left(-\frac{1}{n} l''(\hat{\theta}) + o_p(1)\right)} = \frac{\mathcal{I}(\theta_0)}{\left(-\frac{1}{n} l''(\hat{\theta}) + o_p(1)\right)} \\
&\xrightarrow{\mathcal{L}} 1
\end{align*}$$

More details. Notice that $-\frac{1}{n} l''(\hat{\theta}) = -\frac{1}{n} \sum_{i=1}^{n} \frac{\partial^2}{\partial \theta^2} \log p_{\theta}(X_i)|_{\theta = \hat{\theta}}$, and $\hat{\theta} = \hat{\theta}(X_1, ..., X_n)$ is a function of the data, such that the addends in the sum are not independent and we cannot directly apply Theorem 13 of Lecture notes 4, which is a version of WLLN requiring independence. By Taylor’s expansion:

$$l''(\hat{\theta}) = l''(\theta_0) + l''(\hat{\theta})(\hat{\theta} - \theta_0), \quad \text{with } \hat{\theta} = \theta_0 + \tau(\hat{\theta} - \theta_0), \tau \in (0, 1)$$
such that, for reasons explained earlier,

\[-\frac{1}{n} l''(\hat{\theta}) = -\frac{1}{n} l''(\theta_0) + o_p(1/n)\]

Now, \(-\frac{1}{n} l''(\theta_0)\) is a sum of independent addends whose finite expectation is \(I(\theta_0)\), such that by Theorem 13 of Lecture notes 4 we have \(-\frac{1}{n} l''(\theta_0) \xrightarrow{P} I(\theta_0)\). Finally, by Slutsky’s Theorem

\[-\frac{1}{n} l''(\hat{\theta}) = -\frac{1}{n} l''(\theta_0) + o_p(1/n) \xrightarrow{P} I(\theta_0)\]

**Problem 4.** [20 points]

(a) This is a particular case of Example 3 in Lecture Notes 10. \(W = \bar{Y}_n \frac{1}{\sqrt{n}}\) has distribution \(N(0,1)\) under \(\theta = 0\). Then:

\[P_0 (W > z_\alpha) = \alpha\]

Thus \(c_n = \frac{z_\alpha}{\sqrt{n}}\)

(b) For the previous test with critical value \(c_n\), at an arbitrary alternative \(\theta\) we have:

\[\beta(\theta) = P_\theta (\bar{Y}_n > c_n) = P_\theta \left( \frac{\bar{Y}_n - \theta}{1/\sqrt{n}} > z_\alpha - \theta \sqrt{n} \right) = 1 - \Phi \left( z_\alpha - \theta \sqrt{n} \right)\]

So in particular at a local alternative of the form \(\theta_n = a/\sqrt{n}\) the power is given by:

\[\beta(\theta_n) = 1 - \Phi \left( z_\alpha - a \right)\]

which is constant.