(5) (a) **Delta method**

Let $X_1, \dots, X_n$ be i.i.d. from normal distribution with unknown mean μ and known variance σ^2. Let $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$. Find the limiting distribution of $\sqrt{n}\left(\bar{X}^3 - c\right)$ for an appropriate constant c.

(b) **Logit transformation**

Let $X_n \sim bin(n, p)$. Consider the logit transformation, defined by

$$logit(y) = \ln \frac{y}{1-y}, \qquad 0 < y < 1.$$

Determine the approximate distribution of $logit\left(\frac{X_n}{n}\right)$.

a) By the CLT, $\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{d} Y \sim \mathcal{N}(0,1)$

$Y_n := \frac{\bar{X}_n}{\sigma}$, $\theta := \frac{\mu}{\sigma}$, then $\sqrt{n}(Y_n - \theta) \xrightarrow{d} Y \sim \mathcal{N}(0,1)$

$g: \mathbb{R} \to \mathbb{R}: y \mapsto (\sigma y)^3$, $\Rightarrow g'(y) = 3\sigma^3 y^2$, and by the delta method,

$\sqrt{n}\left(\bar{X}_n^3 - \mu^3\right) = \sqrt{n}\left((\sigma Y_n)^3 - (\sigma\theta)^3\right) = \sqrt{n}\left(g(Y_n) - g(\theta)\right) \xrightarrow{d} g'(\theta) Y = 3\sigma^3\theta^2 Y$

and $3\sigma^3\theta^2 Y = 3\sigma^3 \frac{\mu^2}{\sigma^2} Y = 3\sigma\mu^2 Y \sim \mathcal{N}\left(0, (3\sigma\mu^2)^2\right)$

b) Let $Y_1, \dots, Y_n \sim bin(1,p)$; $X_n = \sum_{i=1}^{n} Y_i \sim bin(n,p)$, $\mathbb{E}(X_n) = np$, $\mathbb{V}ar(X_n) = np(1-p)$

$\mathbb{E}(Y_i) = p$, $\mathbb{V}ar(Y_i) = p(1-p)$

•) For $p \notin \{0,1\}$

By CLT: $\sqrt{n}\,\frac{\frac{X_n}{n} - p}{\sqrt{p(1-p)}} = \frac{\sqrt{n}(X_n - np)}{n\sqrt{p(1-p)}} = \frac{X_n - np}{\sqrt{np(1-p)}} \xrightarrow{d} Z \sim \mathcal{N}(0,1)$

We define $Z_n := \frac{X_n}{n\sqrt{p(1-p)}}$ and $\theta := \sqrt{\frac{p}{1-p}}$, then

$\sqrt{n}(Z_n - \theta) \xrightarrow{d} \mathcal{N}(0,1)$, we define $g: (0,1) \to \mathbb{R}: z \mapsto logit\left(\sqrt{p(1-p)}\,z\right)$, hence

$g'(z) = \frac{1 - \sqrt{p(1-p)}\,z}{z}\left(\frac{\sqrt{p(1-p)}}{1 - \sqrt{p(1-p)}\,z} + z\sqrt{p(1-p)}\left(1 - \sqrt{p(1-p)}\,z\right)^{-2}\sqrt{p(1-p)}\right)$

$= \frac{1 - \sqrt{p(1-p)}\,z}{\sqrt{p(1-p)}\,z} \cdot \frac{\sqrt{p(1-p)}\left(1 - \sqrt{p(1-p)}\,z\right) + z p(1-p)}{\left(1 - \sqrt{p(1-p)}\,z\right)^2} = \frac{1}{z\left(1 - \sqrt{p(1-p)}\,z\right)}$ and $g(\theta) = \log\left(\frac{p}{1-p}\right)$

by application of the delta method we obtain and

$\sqrt{n}\left(logit\left(\frac{X_n}{n}\right) - logit(p)\right) = \sqrt{n}\left(g(Z_n) - g(\theta)\right) \xrightarrow{d} g'(\theta) Z = \frac{\sqrt{1-p}}{\sqrt{p}(1-p)} Z = \frac{1}{\sqrt{p(1-p)}} Z \sim \mathcal{N}\left(0, \frac{1}{p(1-p)}\right)$

hence, $logit\left(\frac{X_n}{n}\right) \approx \mathcal{N}\left(logit(p), \frac{1}{np(1-p)}\right)$