1. Suppose $Y_1, Y_2, ..., Y_n$ is an iid sample from a Poisson population distribution with mean $\lambda > 0$. The Central Limit Theorem (Chapter 7) assures us that

$$\frac{\overline{Y} - \lambda}{\sqrt{\frac{\lambda}{n}}} \xrightarrow{d} \mathcal{N}(0,1), \quad \text{as } n \to \infty.$$

(a) Carefully argue that

$$\frac{\overline{Y} - \lambda}{\sqrt{\frac{\overline{Y}}{n}}} \quad \text{and} \quad \frac{\overline{Y} - \lambda}{\sqrt{\frac{S^2}{n}}}$$

also converge in distribution to $\mathcal{N}(0,1)$, as $n \to \infty$. *Hint:* Argue that $\overline{Y}$ and S^2 are both consistent estimators of λ and then use Slutsky's Theorem.
(b) Derive large-sample $100(1-\alpha)\%$ confidence intervals for λ using each quantity in part (a). Note that each one is a large-sample pivot.
(c) Suppose $n = 100$, $\lambda = 10$, and you calculate

$$\frac{\overline{Y} - 10}{\sqrt{\frac{\overline{Y}}{100}}}.$$

Between what two values would you expect this statistic to fall with probability approximately equal to 0.95? What might you conclude if it fell well outside this range?

2. Suppose $Y_1, Y_2, ..., Y_n$ is an iid sample from a Bernoulli(p) population distribution, where $0 < p < 1$. The population pmf is

$$p_Y(y|p) = \begin{cases} p^y(1-p)^{1-y}, & y = 0, 1 \\ 0, & \text{otherwise.} \end{cases}$$

(a) Prove that

$$\widehat{p} = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

is the maximum likelihood estimator (MLE) of p.
(b) Use the large-sample properties of MLEs to show that

$$\widehat{p} \sim \mathcal{AN}\left(p, \ \frac{p(1-p)}{n}\right),$$

for large n. Note this is the same result you would get by applying the CLT directly.

Delta Method. Suppose $Y_1, Y_2, ..., Y_n$ is an iid sample from a population distribution described by $p_Y(y|\theta)$ or $f_Y(y|\theta)$, and suppose $\widehat{\theta}$ is the MLE for θ. Under certain regularity conditions, we know

$$\frac{\widehat{\theta}-\theta}{\sqrt{\frac{v(\theta)}{n}}} \xrightarrow{d} \mathcal{N}(0,1),$$

as $n \to \infty$, where

$$v(\theta) = \frac{1}{E\left[-\frac{\partial^2}{\partial\theta^2}\ln p_Y(Y|\theta)\right]} \quad \text{(discrete case)}$$

$$v(\theta) = \frac{1}{E\left[-\frac{\partial^2}{\partial\theta^2}\ln f_Y(Y|\theta)\right]} \quad \text{(continuous case)}.$$

This means MLEs are asymptotically normal. *The Delta Method says that functions of MLEs are also asymptotically normal.* Specifically, suppose g is a continuous and differentiable function. Then,

$$\frac{g(\widehat{\theta})-g(\theta)}{\sqrt{\frac{[g'(\theta)]^2 v(\theta)}{n}}} \xrightarrow{d} \mathcal{N}(0,1),$$

as $n \to \infty$, where $g'(\theta) = \partial g(\theta)/\partial\theta$. To see why this is true, let's sketch a proof. First, write

$$g(\widehat{\theta}) = g(\theta) + g'(\theta)(\widehat{\theta}-\theta) + \text{higher order terms}$$

using a Taylor series expansion. After rearranging and multiplying by $\sqrt{n}$, we have

$$\sqrt{n}[g(\widehat{\theta}) - g(\theta)] \approx g'(\theta)\sqrt{n}(\widehat{\theta}-\theta).$$

The term $\sqrt{n}(\widehat{\theta}-\theta)$ converges in distribution to $\mathcal{N}(0, v(\theta))$ by assumption and hence $g'(\theta)\sqrt{n}(\widehat{\theta}-\theta)$ converges in distribution to

$$\mathcal{N}\left(0, [g'(\theta)]^2 v(\theta)\right).$$

Because all higher order terms (even when multiplied by $\sqrt{n}$) converge in probability to zero, Slutsky's Theorem gives

$$\sqrt{n}[g(\widehat{\theta}) - g(\theta)] \xrightarrow{d} \mathcal{N}\left(0, [g'(\theta)]^2 v(\theta)\right) \iff \frac{g(\widehat{\theta})-g(\theta)}{\sqrt{\frac{[g'(\theta)]^2 v(\theta)}{n}}} \xrightarrow{d} \mathcal{N}(0,1),$$

or, in other words,

$$g(\widehat{\theta}) \sim \mathcal{AN}\left(g(\theta),\ \frac{[g'(\theta)]^2 v(\theta)}{n}\right),$$

for large n.

3. In Problem 2, use the Delta Method to show

$$\frac{\widehat{p}}{1-\widehat{p}} \sim \mathcal{AN}\left(\frac{p}{1-p}, \ \frac{p}{n(1-p)^3}\right),$$

for large n. The statistic

$$\frac{\widehat{p}}{1-\widehat{p}}$$

is called the *sample odds. Hint:* Work with $g(p) = p/(1-p)$.

4. Suppose $Y_1, Y_2, ..., Y_n$ is an iid sample from an exponential distribution with mean $\beta > 0$.
(a) Find the MLE of β and derive its large-sample distribution.
(b) Find the MLE of β^2 and derive its large-sample distribution.
(c) Find the MLE of $P(Y > 1) = 1 - e^{-1/\beta}$ and derive its large-sample distribution.

5. Suppose $X_1, X_2, ..., X_n$ is an iid $\mathcal{N}(\mu, c^2\mu^2)$ sample, where c^2 is known. Let $\widetilde{\mu}$ and $\widehat{\mu}$ denote the method of moments and maximum likelihood estimators of μ, respectively.
(a) Show that

$$\widetilde{\mu} = \overline{X} \qquad \text{and} \qquad \widehat{\mu} = \frac{\sqrt{\overline{X}^2 + 4c^2 m_2'} - \overline{X}}{2c^2},$$

where $m_2' = n^{-1}\sum_{i=1}^n X_i^2$ is the second sample moment.
(b) Prove that both estimators $\widetilde{\mu}$ and $\widehat{\mu}$ are consistent.
(c) Show that

$$\widetilde{\mu} \sim \mathcal{AN}\left(\mu, \frac{\sigma^2_{\widetilde{\mu}}}{n}\right) \qquad \text{and} \qquad \widehat{\mu} \sim \mathcal{AN}\left(\mu, \frac{\sigma^2_{\widehat{\mu}}}{n}\right)$$

for large n and calculate $\sigma^2_{\widetilde{\mu}}$ and $\sigma^2_{\widehat{\mu}}$. Which estimator is more efficient?