Large Sample Theory
Ferguson

Exercises, Section 7, Functions of the Sample Moments.

1. Let $X_1, X_2, \ldots$ be i.i.d. random variables with mean μ and variance σ^2. Find the asymptotic distribution of $R_n = \sum_{i=1}^n X_{2i-1}/\sum_{i=1}^n X_{2i}$ for (a) $\mu \neq 0$, and for (b) $\mu = 0$.

2. Professor Bliss has at hand a large sample $X_1, \ldots, X_n$, from the double exponential distribution with density $f(x) = (1/(2\tau))e^{-|x-\mu|/\tau}$, having mean μ and mean deviation $\mathrm{E}|X - \mu| = \tau$. He knows enough to estimate μ by the sample median, m_n, and he knows he should use $(1/n)\sum |X_i - m_n|$ to estimate the mean deviation (these are the MLE's), or $(1/n)\sum(X_i - m_n)^2$ to estimate the variance, $\sigma^2 = 2\tau^2$, but he doesn't quite know what the sampling distribution might be. He decides instead to use the sample variance, $(1/n)\sum(X_i - \overline{X}_n)^2$, to estimate σ^2, and to get confidence intervals for σ^2 using the chi-square tables. How well is Professor Bliss doing in his confidence intervals for σ^2? (You may assume n large.)

3. Let X have the Poisson distribution, $\mathcal{P}(\lambda)$. We know that $(X - \lambda)/\sqrt{\lambda} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$ as $\lambda \to \infty$, and we say $X \sim \mathcal{N}(\lambda, \lambda)$ for large λ.
(a) Show $\log(X) \sim \mathcal{N}(\log(\lambda), \lambda^{-1})$ for large λ.
(b) Show $X^2 \sim \mathcal{N}(\lambda^2, 4\lambda^3)$ for large λ.
(c) Is it true that $e^X \sim \mathcal{N}(e^\lambda, \text{something})$ for large λ?

4. Let $X_1, \ldots, X_n$ be a sample from the geometric distribution with mass function, $\mathrm{P}(X = x) = (1 - \theta)\theta^x$ for $x = 0, 1, \ldots$, where $0 < \theta < 1$ is a success probability. Let $S_n = \sum_1^n X_i$ denote the total number of successes, and $T_n = \sum_1^n \mathrm{I}(X_i > 0)$ denote the number of trials that had at least one success.
(a) Find the joint asymptotic distribution of (S_n, T_n).
(b) Find the joint asymptotic distribution of (U_n, V_n), where $U_n = S_n/T_n$ and $V_n = n - T_n$.

5. To estimate a parameter, θ^2, you are given the choice of the following two possibilities: (1) the estimate $\overline{X}_n^2$, based on a sample, $X_1, \ldots, X_n$ from the gamma distribution, $\mathcal{G}(\theta, 1)$, and (2) the estimate $\overline{Y}_n$, based on a sample, $Y_1, \ldots, Y_n$ from the gamma distribution, $\mathcal{G}(\theta^2, 1)$. If n is large, which would you choose? (The answer depends on θ.)

6. If $\sqrt{n}(\overline{X}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2)$ as $n \to \infty$, what is the asymptotic distribution of $|\overline{X}_n|$? (Consider the cases $\theta = 0$ and $\theta \neq 0$ separately.)

7. Let $X_1, \ldots, X_n$ be a sample from $\mathcal{N}(\theta, \sigma^2)$ with σ^2 known. For a fixed number a, let $p = \mathrm{P}(X_i > a) = 1 - \Phi((a - \theta)/\sigma) = \Phi((\theta - a)/\sigma)$. The maximum likelihood estimate of p is therefore $\hat{p}_n = \Phi((\overline{X}_n - a)/\sigma)$. Find the asymptotic distribution of $\sqrt{n}(\hat{p}_n - p)$.

8. Let $X_1, \ldots, X_n$ be i.i.d. with mean zero and positive finite sixth moment. Let $\mu_k = \mathrm{E}(X^k)$ denote the population moments and $m_k = (1/n)\sum_1^n X_i^k$ denote the sample moments. Then m_2 is a reasonable estimate of μ_2 and has asymptotic distribution

$$\sqrt{n}(m_2 - \mu_2) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \mu_4 - \mu_2^2).$$

Show that the estimate of μ_2 given by

$$\hat{\sigma}^2 = m_2 - \frac{m_1 m_3}{m_2}$$

has an asymptotic normal distribution,

$$\sqrt{n}(\hat{\sigma}^2 - \mu_2) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \tau^2).$$

with some asymptotic variance τ^2. Find τ^2 and show that $\tau^2 \le \mu_4 - \mu_2^2$, with equality if and only if $\mu_3 = 0$. Note that for two-point distributions, $\tau^2 = 0$.

9. (a) Suppose $\sqrt{n}(Z_n - \sigma^2) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 2\sigma^4)$, where $\sigma > 0$. Find the asymptotic distribution of $\sqrt{n}(\sqrt{Z_n} - \sigma)$.

(b) Find the approximation given by the second order Taylor expansion to the asymptotic distribution of $\sqrt{n}(\sqrt{Z_n} - \sigma)$.

(c) Take $n = 10$, $\sigma = 1$ and suppose the original distribution of $\sqrt{n}(Z_n - \sigma^2)$ is exactly normal. Find the exact probability, $\mathrm{P}(\sqrt{n}(\sqrt{Z_n} - \sigma) > .5)$, and compare it to the approximations given by (a) and (b).

(d) Suppose the distribution of Z_n is not normal but instead that nZ_n/σ^2 is exactly χ^2_n, as it would be if Z_n were the sample variance of a sample of size $n+1$ from a normal distribution with variance σ^2 (i.e. $Z_n = (1/n)\sum_1^{n+1}(X_i - \overline{X}_{n+1})^2 = s_x^2$). Now find the exact probability $\mathrm{P}(\sqrt{n}(\sqrt{Z_n} - \sigma) > .5)$ for $n = 10$ and $\sigma = 1$, and compare it to the approximations given by (a) using the Edgeworth expansions. (Note that Table 1 has fortuitously been constructed for the normalized χ^2_{10} distribution.)

10. For convenience, Cramér's Theorem has been stated assuming $g'(x)$ is continuous in a neighborhood of μ. It also holds under the weaker assumption that $g'(x)$ exists at μ in the sense that $\dfrac{g(x) - g(\mu)}{x - \mu} \to g'(\mu)$ as $x \to \mu$, $x \ne \mu$. Show this in one dimension:

Theorem. *Let $g(x)$ be defined in a neighborhood of μ and assume that $g'(x)$ exists at μ. If $b_n(X_n - \mu) \xrightarrow{\mathcal{L}} X$, where b_n is a sequence of numbers tending to infinity. Then $b_n(g(X_n) - g(\mu)) \xrightarrow{\mathcal{L}} g'(\mu)X$.*