Lecture 28: Asymptotic inference

Statistical inference based on asymptotic criteria and approximations is called *asymptotic statistical inference* or simply *asymptotic inference*.
We have previously considered asymptotic estimation.
We now focus on asymptotic hypothesis tests and confidence sets.

Hypothesis tests

Definition 2.13. Let $X = (X_1, ..., X_n)$ be a sample from $P \in \mathcal{P}$ and $T_n(X)$ be a test for $H_0 : P \in \mathcal{P}_0$ versus $H_1 : P \in \mathcal{P}_1$.
(i) If $\limsup_n \alpha_{T_n}(P) \le \alpha$ for any $P \in \mathcal{P}_0$, then α is an *asymptotic significance level* of T_n.
(ii) If $\lim_{n\to\infty} \sup_{P\in\mathcal{P}_0} \alpha_{T_n}(P)$ exists, then it is called the *limiting size* of T_n.
(iii) T_n is called *consistent* if and only if the type II error probability converges to 0, i.e., $\lim_{n\to\infty}[1 - \alpha_{T_n}(P)] = 0$, for any $P \in \mathcal{P}_1$.
(iv) T_n is called *Chernoff-consistent* if and only if T_n is consistent *and* the type I error probability converges to 0, i.e., $\lim_{n\to\infty} \alpha_{T_n}(P) = 0$, for any $P \in \mathcal{P}_0$. T_n is called *strongly Chernoff-consistent* if and only if T_n is consistent and the limiting size of T_n is 0.

Obviously if T_n has size (or significance level) α for all n, then its limiting size (or asymptotic significance level) is α.
If the limiting size of T_n is $\alpha \in (0, 1)$, then for any $\epsilon > 0$, T_n has size $\alpha + \epsilon$ for all $n \ge n_0$, where n_0 is independent of P.
Hence T_n has level of significance $\alpha + \epsilon$ for any $n \ge n_0$.
However, if $\mathcal{P}_0$ is not a parametric family, it is likely that the limiting size of T_n is 1 (see, e.g., Example 2.37).
This is the reason why we consider the weaker requirement in Definition 2.13(i).
If T_n has asymptotic significance level α, then for any $\epsilon > 0$, $\alpha_{T_n}(P) < \alpha+\epsilon$ for all $n \ge n_0(P)$ but $n_0(P)$ depends on $P \in \mathcal{P}_0$; and there is no guarantee that T_n has significance level $\alpha + \epsilon$ for any n.

The consistency in Definition 2.13(iii) only requires that the type II error probability converge to 0.
We may define uniform consistency to be $\lim_{n\to\infty} \sup_{P\in\mathcal{P}_1}[1 - \alpha_{T_n}(P)] = 0$, but it is not satisfied in most problems.
If $\alpha \in (0, 1)$ is a pre-assigned level of significance for the problem, then a consistent test T_n having asymptotic significance level α is called *asymptotically correct*, and a consistent test having limiting size α is called *strongly* asymptotically correct.

The Chernoff-consistency (or strong Chernoff-consistency) in Definition 2.13(iv) requires that both types of error probabilities converge to 0.
Mathematically, Chernoff-consistency (or strong Chernoff-consistency) is better than asymptotic correctness (or strongly asymptotic correctness).
After all, both types of error probabilities should decrease to 0 if sampling can be continued indefinitely.
However, if α is chosen to be small enough so that error probabilities smaller than α can

be practically treated as 0, then the asymptotic correctness (or strongly asymptotic correctness) is enough, and is probably preferred, since requiring an unnecessarily small type I error probability usually results in an unnecessary increase in the type II error probability.

Example 2.37. Consider the testing problem $H_0 : \mu \leq \mu_0$ versus $H_1 : \mu > \mu_0$ based on i.i.d. $X_1, ..., X_n$ with $EX_1 = \mu \in \mathcal{R}$. If each X_i has the $N(\mu, \sigma^2)$ distribution with a known σ^2, then the test $T_{c_\alpha} I_{(c_\alpha, \infty)}(\bar{X})$ with $c_\alpha = \sigma z_{1-\alpha}/\sqrt{n} + \mu_0$ and $\alpha \in (0, 1)$ has size α (and, therefore, limiting size α).
For any $\mu > \mu_0$,

$$1 - \alpha_{T_{c_\alpha}}(\mu) = \Phi\left(z_{1-\alpha} + \frac{\sqrt{n}(\mu_0 - \mu)}{\sigma}\right) \to 0 \tag{1}$$

as $n \to \infty$.
This shows that T_{c_α} is consistent and, hence, is strongly asymptotically correct.
The convergence in (1) is not uniform in $\mu > \mu_0$, but is uniform in $\mu > \mu_1$ for any fixed $\mu_1 > \mu_0$.

Since the size of T_{c_α} is α for all n, T_{c_α} is not Chernoff-consistent.
A strongly Chernoff-consistent test can be obtained as follows.
Let

$$\alpha_n = 1 - \Phi(\sqrt{n} a_n), \tag{2}$$

where a_n's are positive numbers satisfying $a_n \to 0$ and $\sqrt{n} a_n \to \infty$.
Let T_n be T_{c_α} with $\alpha = \alpha_n$ for each n.
Then, T_n has size α_n.
Since $\alpha_n \to 0$, The limiting size of T_n is 0.
On the other hand, (1) still holds with α replaced by α_n.
This follows from the fact that

$$z_{1-\alpha_n} + \frac{\sqrt{n}(\mu_0 - \mu)}{\sigma} = \sqrt{n}\left(a_n + \frac{\mu_0 - \mu}{\sigma}\right) \to -\infty$$

for any $\mu > \mu_0$.
Hence T_n is strongly Chernoff-consistent.
However, if $\alpha_n < \alpha$, then, from the left-hand side of (1), $1 - \alpha_{T_{c_\alpha}}(\mu) < 1 - \alpha_{T_n}(\mu)$ for any $\mu > \mu_0$.

We now consider the case where the population P is not in a parametric family.
We still assume that $\sigma^2 = \mathrm{Var(X_i)}$ is known.
Using the CLT, we can show that for $\mu > \mu_0$,

$$\lim_{n\to\infty}[1 - \alpha_{T_{c_\alpha}}(\mu)] = \lim_{n\to\infty} \Phi\left(z_{1-\alpha} + \frac{\sqrt{n}(\mu_0 - \mu)}{\sigma}\right) = 0,$$

i.e., T_{c_α} is still consistent.
For $\mu \leq \mu_0$,

$$\lim_{n\to\infty} \alpha_{T_{c_\alpha}}(\mu) = 1 - \lim_{n\to\infty} \Phi\left(z_{1-\alpha} + \frac{\sqrt{n}(\mu_0 - \mu)}{\sigma}\right),$$

which equals α if $\mu = \mu_0$ and 0 if $\mu < \mu_0$.
Thus, the asymptotic significance level of T_{c_α} is α.
Combining these two results, we know that T_{c_α} is asymptotically correct.
However, if $\mathcal{P}$ contains all possible populations on $\mathcal{R}$ with finite second moments, then one can show that the limiting size of T_{c_α} is 1 (exercise).
For α_n defined by (2), we can show that $T_n = T_{c_\alpha}$ with $\alpha = \alpha_n$ is Chernoff-consistent (exercise).
But T_n is not strongly Chernoff-consistent if $\mathcal{P}$ contains all possible populations on $\mathcal{R}$ with finite second moments.

Example. Let $(X_1, ..., X_n)$ be a random sample from the exponential distribution $E(0, \theta)$, where $\theta \in (0, \infty)$.
Consider the hypotheses $H_0 : \theta \leq \theta_0$ versus $H_1 : \theta > \theta_0$, where $\theta_0 > 0$ is a fixed constant.
Let $T_c = I_{(c,\infty)}(\bar{X})$, where $\bar{X}$ is the sample mean.
$\bar{X}/\theta$ has the gamma distribution with shape parameter n and scale parameter θ/n.
Let $G_{n,\theta}$ denote the cumulative distribution function of this distribution and $c_{n,\alpha}$ be the constant satisfying $G_{n,\theta_0}(c_{n,\alpha}) = 1 - \alpha$.
Then,

$$\sup_{\theta \leq \theta_0} P(T_{c_{n,\alpha}} = 1) = \sup_{\theta \leq \theta_0} [1 - G_{n,\theta}(c_{n,\alpha})] = 1 - G_{n,\theta_0}(c_{n,\alpha}) = \alpha,$$

i.e., the size of $T_{c_{n,\alpha}}$ is α.
Since the power of $T_{c_{n,\alpha}}$ is $P(T_{c_{n,\alpha}} = 1) = P(\bar{X} > c_{n,\alpha})$ for $\theta > \theta_0$ and, by the law of large numbers, $\bar{X} \to_p \theta$, the consistency of $T_{c_{n,\alpha}}$ follows if we can show that $\lim_{n\to\infty} c_{n,\alpha} = \theta_0$.
By the central limit theorem, $\sqrt{n}(\bar{X} - \theta) \to_d N(0, \theta^2)$.
Hence, $\sqrt{n}(\frac{\bar{X}}{\theta} - 1) \to_d N(0, 1)$.
By Pólya's theorem (Proposition 1.16),

$$\lim_{n\to\infty} \sup_t \left| P\Big(\sqrt{n}\left(\tfrac{\bar{X}}{\theta} - 1\right) \leq t\Big) - \Phi(t) \right| = 0,$$

where Φ is the cumulative distribution function of the standard normal distribution.
When $\theta = \theta_0$,

$$\alpha = P(\bar{X} \geq c_{n,\alpha}) = P\Big(\sqrt{n}\left(\tfrac{\bar{X}}{\theta_0} - 1\right) \geq \sqrt{n}\left(\tfrac{c_{n,\alpha}}{\theta_0} - 1\right)\Big).$$

Hence

$$\lim_{n\to\infty} \Phi\Big(\sqrt{n}\left(\tfrac{c_{n,\alpha}}{\theta_0} - 1\right)\Big) = 1 - \alpha,$$

which implies $\lim_{n\to\infty} \sqrt{n}(\frac{c_{n,\alpha}}{\theta_0} - 1) = \Phi^{-1}(1 - \alpha)$ and, thus, $\lim_{n\to\infty} c_{n,\alpha} = \theta_0$.

Let $\{a_n\}$ be a sequence of positive numbers such that $\lim_{n\to\infty} a_n = 0$ and $\lim_{n\to\infty} \sqrt{n}a_n = \infty$.
Let $\alpha_n = 1 - \Phi(\sqrt{n}a_n)$ and $b_n = c_{n,\alpha_n}$.
From the previous derivation, the size of T_{b_n} is α_n, which converges to 0 as $n \to \infty$ since $\lim_{n\to\infty} \sqrt{n}a_n = \infty$.
Using the previous argument, we can show that

$$\lim_{n\to\infty} \left| 1 - \alpha_n - \Phi\Big(\sqrt{n}\left(\tfrac{c_{n,\alpha_n}}{\theta_0} - 1\right)\Big) \right| = 0,$$

which implies that

$$\lim_{n\to\infty} \frac{\sqrt{n}}{\Phi^{-1}(1-\alpha_n)} \left(\tfrac{c_{n,\alpha_n}}{\theta_0} - 1\right) = 1.$$

Since $1 - \alpha_n = \Phi(\sqrt{n} a_n)$, this implies that $\lim_{n\to\infty} c_{n,\alpha_n} = \theta_0$.
Since $b_n = c_{n,\alpha_n}$, the test T_{b_n} is Chernoff-consistent.

Confidence sets

Definition 2.14. Let $X = (X_1, ..., X_n)$ be a sample from $P \in \mathcal{P}$, ϑ be a k-vector of parameters related to P, and $C(X)$ be a confidence set for ϑ.
(i) If $\liminf_n P(\vartheta \in C(X)) \geq 1 - \alpha$ for any $P \in \mathcal{P}$, then $1 - \alpha$ is an *asymptotic significance level* of $C(X)$.
(ii) If $\lim_{n\to\infty} \inf_{P\in\mathcal{P}} P(\vartheta \in C(X))$ exists, then it is called the *limiting confidence coefficient* of $C(X)$.

Note that the asymptotic significance level and limiting confidence coefficient of a confidence set are very similar to the asymptotic significance level and limiting size of a test, respectively. Some conclusions are also similar.
For example, in a parametric problem one can often find a confidence set having limiting confidence coefficient $1-\alpha \in (0, 1)$, which implies that for any $\epsilon > 0$, the confidence coefficient of $C(X)$ is $1-\alpha-\epsilon$ for all $n \geq n_0$, where n_0 is independent of P. In a nonparametric problem the limiting confidence coefficient of $C(X)$ might be 0, whereas $C(X)$ may have asymptotic significance level $1 - \alpha \in (0, 1)$, but for any fixed n, the confidence coefficient of $C(X)$ might be 0.