We can apply the notions studied in Examples 4.4.4 and 4.4.5 to give confidence regions for scalar or vector parameters in nonparametric models.

Example 4.4.7. *A Lower Confidence Bound for the Mean of a Nonnegative Random Variable.* Suppose $X_1, \dots, X_n$ are i.i.d. as X and that X has a density $f(t) = F'(t)$, which is zero for $t < 0$ and nonzero for $t > 0$. By integration by parts, if $\mu = \mu(F) = \int_0^\infty tf(t)dt$ exists, then

$$\mu = \int_0^\infty [1 - F(t)]dt.$$

Let $\widehat{F}^-(t)$ and $\widehat{F}^+(t)$ be the lower and upper simultaneous confidence boundaries of Example 4.4.6. Then a $(1-\alpha)$ lower confidence bound for μ is $\underline{\mu}$ given by

$$\underline{\mu} = \int_0^\infty [1 - \widehat{F}^+(t)]dt = \sum_{i \le n(1-d_\alpha)} \left[1 - \left(\frac{i}{n} + d_\alpha\right)\right] [x_{(i+1)} - x_{(i)}] \tag{4.4.9}$$

because for $C(\mathbf{X})$ as in Example 4.4.6, $\underline{\mu} = \inf\{\mu(F) : F \in C(\mathbf{X})\} = \mu(\widehat{F}^+)$ and $\sup\{\mu(F) : F \in C(\mathbf{X})\} = \mu(\widehat{F}^-) = \infty$—see Problem 4.4.19.

Intervals for the case F supported on an interval (see Problem 4.4.18) arise in accounting practice (see Bickel, 1992) where such bounds are discussed and shown to be asymptotically strictly conservative. □

Summary. We define lower and upper confidence bounds (LCBs and UCBs), confidence intervals, and more generally confidence regions. In a parametric model $\{P_\theta : \theta \in \Theta\}$, a level $1-\alpha$ confidence region for a parameter $q(\theta)$ is a set $C(x)$ depending only on the data x such that the probability under P_θ that $C(X)$ covers $q(\theta)$ is at least $1-\alpha$ for all $\theta \in \Theta$. For a nonparametric class $\mathcal{P} = \{P\}$ and parameter $\nu = \nu(P)$, we similarly require $P(C(X) \supset \nu) \ge 1-\alpha$ for all $P \in \mathcal{P}$. We derive the (Student) t interval for μ in the $\mathcal{N}(\mu, \sigma^2)$ model with σ^2 unknown, and we derive an exact confidence interval for the binomial parameter. In a nonparametric setting we derive a simultaneous confidence interval for the distribution function $F(t)$ and the mean of a positive variable X.

4.5 THE DUALITY BETWEEN CONFIDENCE REGIONS AND TESTS

Confidence regions are random subsets of the parameter space that contain the true parameter with probability at least $1-\alpha$. Acceptance regions of statistical tests are, for a given hypothesis H, subsets of the sample space with probability of accepting H at least $1-\alpha$ when H is true. We shall establish a duality between confidence regions and acceptance regions for families of hypotheses.

We begin by illustrating the duality in the following example.

Example 4.5.1. *Two-Sided Tests for the Mean of a Normal Distribution.* Suppose that an established theory postulates the value μ_0 for a certain physical constant. A scientist has reasons to believe that the theory is incorrect and measures the constant n times obtaining measurements $X_1, \ldots, X_n$. Knowledge of his instruments leads him to assume that the X_i are independent and identically distributed normal random variables with mean μ and variance σ^2. If any value of μ other than μ_0 is a possible alternative, then it is reasonable to formulate the problem as that of testing $H : \mu = \mu_0$ versus $K : \mu \neq \mu_0$.

We can base a size α test on the level $(1-\alpha)$ confidence interval (4.4.1) we constructed for μ as follows. We accept H, if and only if, the postulated value μ_0 is a member of the level $(1-\alpha)$ confidence interval

$$[\bar{X} - st_{n-1}\left(1 - \tfrac{1}{2}\alpha\right)/\sqrt{n},\ \bar{X} + st_{n-1}\left(1 - \tfrac{1}{2}\alpha\right)/\sqrt{n}]. \tag{4.5.1}$$

If we let $T = \sqrt{n}(\bar{X} - \mu_0)/s$, then our test accepts H, if and only if, $-t_{n-1}\left(1 - \frac{1}{2}\alpha\right) \leq T \leq t_{n-1}\left(1 - \frac{1}{2}\alpha\right)$. Because $P_\mu[|T| = t_{n-1}\left(1 - \frac{1}{2}\alpha\right)] = 0$ the test is equivalently characterized by rejecting H when $|T| \geq t_{n-1}\left(1 - \frac{1}{2}\alpha\right)$. This test is called *two-sided* because it rejects for both large and small values of the statistic T. In contrast to the tests of Example 4.1.4, it has power against parameter values on either side of μ_0.

Because the same interval (4.5.1) is used for every μ_0 we see that we have, in fact, generated a family of level α tests $\{\delta(\mathbf{X}, \mu)\}$ where

$$\begin{aligned} \delta(\mathbf{X}, \mu) &= 1 \text{ if } \sqrt{n}\tfrac{|\bar{X}-\mu|}{s} \geq t_{n-1}\left(1 - \tfrac{1}{2}\alpha\right) \\ &= 0 \text{ otherwise.} \end{aligned} \tag{4.5.2}$$

These tests correspond to different hypotheses, $\delta(\mathbf{X}, \mu_0)$ being of size α only for the hypothesis $H : \mu = \mu_0$.

Conversely, by starting with the test (4.5.2) we obtain the confidence interval (4.5.1) by finding the set of μ where $\delta(\mathbf{X}, \mu) = 0$.

We achieve a similar effect, generating a family of level α tests, if we start out with (say) the level $(1-\alpha)$ LCB $\bar{X} - t_{n-1}(1-\alpha)s/\sqrt{n}$ and define $\delta^*(\mathbf{X}, \mu)$ to equal 1 if, and only if, $\bar{X} - t_{n-1}(1-\alpha)s/\sqrt{n} \geq \mu$. Evidently,

$$P_{\mu_0}[\delta^*(\mathbf{X}, \mu_0) = 1] = P_{\mu_0}\left[\sqrt{n}\frac{(\bar{X} - \mu_0)}{s} \geq t_{n-1}(1-\alpha)\right] = 1 - \alpha.$$

□

These are examples of a general phenomenon. Consider the general framework where the random vector X takes values in the sample space $\mathcal{X} \subset R^q$ and X has distribution $P \in \mathcal{P}$. Let $\nu = \nu(P)$ be a parameter that takes values in the set $\mathcal{N}$. For instance, in Example 4.4.1, $\mu = \mu(P)$ takes values in $\mathcal{N} = (-\infty, \infty)$, in Example 4.4.2, $\sigma^2 = \sigma^2(P)$ takes values in $\mathcal{N} = (0, \infty)$, and in Example 4.4.5, (μ, σ^2) takes values in $\mathcal{N} = (-\infty, \infty) \times (0, \infty)$. For a function space example, consider $\nu(P) = F$, as in Example 4.4.6, where F is the distribution function of X_i. Here an example of $\mathcal{N}$ is the class of all continuous distribution functions. Let $S = S(X)$ be a map from $\mathcal{X}$ to subsets of $\mathcal{N}$, then S is a $(1-\alpha)$ *confidence region* for ν if the probability that $S(X)$ contains ν is at least $(1-\alpha)$, that is

$$P[\nu \in S(X)] \geq 1 - \alpha, \text{ all } P \in \mathcal{P}.$$

Next consider the testing framework where we test the hypothesis $H = H_{\nu_0} : \nu = \nu_0$ for some specified value ν_0. Suppose we have a test $\delta(X, \nu_0)$ with level α. Then the acceptance region

$$A(\nu_0) = \{x : \delta(x, \nu_0) = 0\}$$

is a subset of $\mathcal{X}$ with probability at least $1 - \alpha$. For some specified ν_0, H may be accepted, for other specified ν_0, H may be rejected. Consider the set of ν_0 for which H_{ν_0} is accepted; this is a random set contained in $\mathcal{N}$ with probability at least $1 - \alpha$ of containing the true value of $\nu(P)$ whatever be P. Conversely, if $S(X)$ is a level $1 - \alpha$ confidence region for ν, then the test that accepts H_{ν_0} if and only if ν_0 is in $S(\mathbf{X})$, is a level α test for H_{ν_0}.

Formally, let $\mathcal{P}_{\nu_0} = \{P : \nu(P) = \nu_0 : \nu_0 \in \mathcal{V}\}$. We have the following.

Duality Theorem. *Let* $S(X) = \{\nu_0 \in \mathcal{N} : X \in A(\nu_0)\}$, *then*

$$P[X \in A(\nu_0)] \geq 1 - \alpha \text{ for all } P \in \mathcal{P}_{\nu_0}$$

if and only if $S(X)$ *is a* $1 - \alpha$ *confidence region for* ν.

We next apply the duality theorem to MLR families:

Theorem 4.5.1. *Suppose* $X \sim P_\theta$ *where* $\{P_\theta : \theta \in \Theta\}$ *is MLR in* $T = T(X)$ *and suppose that the distribution function* $F_\theta(t)$ *of* T *under* P_θ *is continuous in each of the variables* t *and* θ *when the other is fixed. If the equation* $F_\theta(t) = 1 - \alpha$ *has a solution* $\underline{\theta}_\alpha(t)$ *in* Θ, *then* $\underline{\theta}_\alpha(T)$ *is a lower confidence bound for* θ *with confidence coefficient* $1 - \alpha$. *Similarly, any solution* $\bar{\theta}_\alpha(T)$ *of* $F_\theta(T) = \alpha$ *with* $\bar{\theta}_\alpha \in \Theta$ *is an upper confidence bound for* θ *with coefficient* $(1 - \alpha)$. *Moreover, if* $\alpha_1 + \alpha_2 < 1$, *then* $[\underline{\theta}_{\alpha_1}, \bar{\theta}_{\alpha_2}]$ *is confidence interval for* θ *with confidence coefficient* $1 - (\alpha_1 + \alpha_2)$.

Proof. By Corollary 4.3.1, the acceptance region of the UMP size α test of $H : \theta = \theta_0$ versus $K : \theta > \theta_0$ can be written

$$A(\theta_0) = \{x : T(x) \leq t_{\theta_0}(1 - \alpha)\}$$

where $t_{\theta_0}(1 - \alpha)$ is the $1 - \alpha$ quantile of F_{θ_0}. By the duality theorem, if

$$s(t) = \{\theta \in \Theta : t \leq t_\theta(1 - \alpha)\},$$

then $S(T)$ is a $1 - \alpha$ confidence region for θ. By applying F_θ to both sides of $t \leq t_\theta(1 - \alpha)$, we find

$$S(t) = \{\theta \in \Theta : F_\theta(t) \leq 1 - \alpha\}.$$

By Theorem 4.3.1, the power function $P_\theta(T \geq t) = 1 - F_\theta(t)$ for a test with critical constant t is increasing in θ. That is, $F_\theta(t)$ is decreasing in θ. It follows that $F_\theta(t) \leq 1 - \alpha$ iff $\theta \geq \underline{\theta}_\alpha(t)$ and $S(t) = [\underline{\theta}_\alpha, \infty)$. The proofs for the upper confidence bound and interval follow by the same type of argument. $\square$

We next give connections between confidence bounds, acceptance regions, and p-values for MLR families: Let t denote the observed value $t = T(x)$ of $T(X)$ for the datum x, let

$\alpha(t, \theta_0)$ denote the p-value for the UMP size α test of $H : \theta = \theta_0$ versus $K : \theta > \theta_0$, and let

$$A^*(\theta) = T(A(\theta)) = \{T(x) : x \in A(\theta)\}.$$

Corollary 4.5.1. *Under the conditions of Theorem* 4.4.1,

$$\begin{aligned} A^*(\theta) &= \{t : \alpha(t, \theta) \geq \alpha\} = (-\infty, t_\theta(1-\alpha)] \\ S(t) &= \{\theta : \alpha(t, \theta) \geq \alpha\} = [\underline{\theta}_\alpha(t), \infty). \end{aligned}$$

Proof. The p-value is

$$\alpha(t, \theta) = P_\theta(T \geq t) = 1 - F_\theta(t).$$

We have seen in the proof of Theorem 4.3.1 that $1 - F_\theta(t)$ is increasing in θ. Because $F_\theta(t)$ is a distribution function, $1 - F_\theta(t)$ is decreasing in t. The result follows. □

In general, let $\alpha(t, \nu_0)$ denote the p-value of a test $\delta(T, \nu_0) = 1[T \geq c]$ of $H : \nu = \nu_0$ based on a statistic $T = T(X)$ with observed value $t = T(x)$. Then the set

$$C = \{(t, \nu) : \alpha(t, \nu) \leq \alpha\} = \{(t, \nu) : \delta(t, \nu) = 0\}$$

gives the pairs (t, ν) where, for the given t, ν will be accepted; and for the given ν, t is in the acceptance region. We call C the set of *compatible* (t, θ) points. In the (t, θ) plane, vertical sections of C are the confidence regions $S(t)$ whereas horizontal sections are the acceptance regions $A^*(\nu) = \{t : \delta(t, \nu) = 0\}$. We illustrate these ideas using the example of testing $H : \mu = \mu_0$ when $X_1, \ldots, X_n$ are i.i.d. $\mathcal{N}(\mu, \sigma^2)$ with σ^2 known. Let $T = \bar{X}$, then

$$C = \left\{(t, \mu) : |t - \mu| \leq \sigma z \left(1 - \tfrac{1}{2}\alpha\right) / \sqrt{n}\right\}.$$

Figure 4.5.1 shows the set C, a confidence region $S(t_0)$, and an acceptance set $A^*(\mu_0)$ for this example.

Example 4.5.2. *Exact Confidence Bounds and Intervals for the Probability of Success in n Binomial Trials.* Let $X_1, \ldots, X_n$ be the indicators of n binomial trials with probability of success θ. For $\alpha \in (0, 1)$, we seek reasonable *exact* level $(1 - \alpha)$ upper and lower confidence bounds and confidence intervals for θ. To find a lower confidence bound for θ our preceding discussion leads us to consider level α tests for $H : \theta \leq \theta_0$, $\theta_0 \in (0, 1)$. We shall use some of the results derived in Example 4.1.3. Let $k(\theta_0, \alpha)$ denote the critical constant of a level $(1 - \alpha)$ test of H. The corresponding level $(1 - \alpha)$ confidence region is given by

$$C(X_1, \ldots, X_n) = \{\theta : S \leq k(\theta, \alpha) - 1\},$$

where $S = \Sigma_{i=1}^n X_i$.

To analyze the structure of the region we need to examine $k(\theta, \alpha)$. We claim that

(i) $k(\theta, \alpha)$ is nondecreasing in θ.

(ii) $k(\theta, \alpha) \to k(\theta_0, \alpha)$ if $\theta \uparrow \theta_0$.

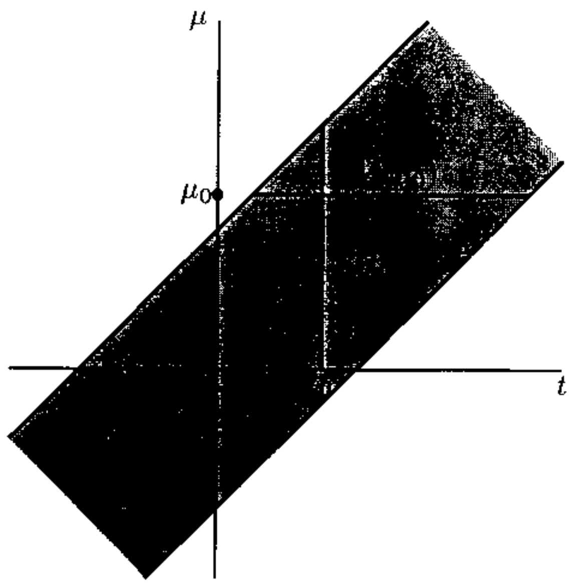

Figure 4.5.1. The shaded region is the compatibility set C for the two-sided test of $H_{\mu_0} : \mu = \mu_0$ in the normal model. $S(t_0)$ is a confidence interval for μ for a given value t_0 of T, whereas $A^*(\mu_0)$ is the acceptance region for H_{μ_0}.

(iii) $k(\theta, \alpha)$ increases by exactly 1 at its points of discontinuity.

(iv) $k(0, \alpha) = 1$ and $k(1, \alpha) = n + 1$.

To prove (i) note that it was shown in Theorem 4.3.1(i) that $P_\theta[S \geq j]$ is nondecreasing in θ for fixed j. Clearly, it is also nonincreasing in j for fixed θ. Therefore, $\theta_1 < \theta_2$ and $k(\theta_1, \alpha) > k(\theta_2, \alpha)$ would imply that

$$\alpha \geq P_{\theta_2}[S \geq k(\theta_2, \alpha)] \geq P_{\theta_2}[S \geq k(\theta_2, \alpha) - 1] \geq P_{\theta_1}[S \geq k(\theta_1, \alpha) - 1] > \alpha,$$

a contradiction.

The assertion (ii) is a consequence of the following remarks. If θ_0 is a discontinuity point o $k(\theta, \alpha)$, let j be the limit of $k(\theta, \alpha)$ as $\theta \uparrow \theta_0$. Then $P_\theta[S \geq j] \leq \alpha$ for all $\theta < \theta_0$ and, hence, $P_{\theta_0}[S \geq j] \leq \alpha$. On the other hand, if $\theta > \theta_0$, $P_\theta[S \geq j] > \alpha$. Therefore, $P_{\theta_0}[S \geq j] = \alpha$ and $j = k(\theta_0, \alpha)$. The claims (iii) and (iv) are left as exercises.

From (i), (ii), (iii), and (iv) we see that, if we define

$$\underline{\theta}(S) = \inf\{\theta : k(\theta, \alpha) = S + 1\},$$

then

$$C(\mathbf{X}) = \begin{cases} (\underline{\theta}(S), 1] & \text{if } S > 0 \\ [0, 1] & \text{if } S = 0 \end{cases}$$

and $\underline{\theta}(S)$ is the desired level $(1-\alpha)$ LCB for θ.[(2)] Figure 4.5.2 portrays the situation. From our discussion, when $S > 0$, then $k(\underline{\theta}(S), \alpha) = S$ and, therefore, we find $\underline{\theta}(S)$ as the unique solution of the equation,

$$\sum_{r=S}^{n} \binom{n}{r} \theta^r (1-\theta)^{n-r} = \alpha.$$

When $S = 0$, $\underline{\theta}(S) = 0$.

Similarly, we define

$$\bar{\theta}(S) = \sup\{\theta : j(\theta, \alpha) = S - 1\}$$

where $j(\theta, \alpha)$ is given by,

$$\sum_{r=0}^{j(\theta,\alpha)} \binom{n}{r} \theta^r (1-\theta)^{n-r} \le \alpha < \sum_{r=0}^{j(\theta,\alpha)+1} \binom{n}{r} \theta^r (1-\theta)^{n-r}.$$

Then $\bar{\theta}(S)$ is a level $(1-\alpha)$ UCB for θ and when $S < n$, $\bar{\theta}(S)$ is the unique solution of

$$\sum_{r=0}^{S} \binom{n}{r} \theta^r (1-\theta)^{n-r} = \alpha.$$

When $S = n$, $\bar{\theta}(S) = 1$. Putting the bounds $\underline{\theta}(S)$, $\bar{\theta}(S)$ together we get the confidence interval $[\underline{\theta}(S), \bar{\theta}(S)]$ of level $(1-2\alpha)$. These intervals can be obtained from computer packages that use algorithms based on the preceding considerations. As might be expected, if n is large, these bounds and intervals differ little from those obtained by the first approximate method in Example 4.4.3.

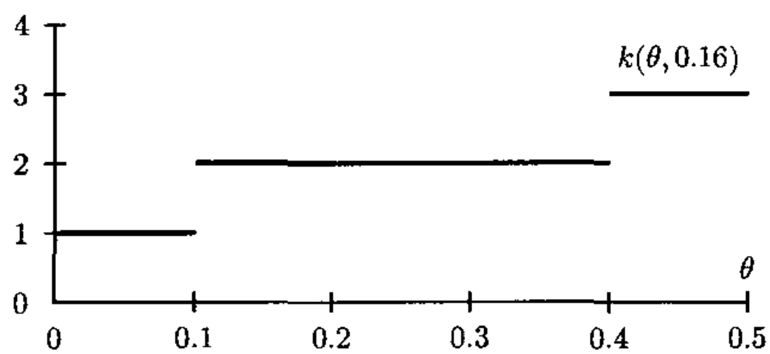

Figure 4.5.2. Plot of $k(\theta, 0.16)$ for $n = 2$.