

1. Introduction

The Basic Statistical Model

As usual, our starting point is a **random experiment** with an underlying **sample space** and a **probability measure** \mathbb{P} . In the basic statistical model, we have an observable **random variable** X taking values in a set S . In general, X can have quite a complicated structure. For example, if the experiment is to sample n objects from a population and record various measurements of interest, then

$$X = (X_1, X_2, \dots, X_n)$$

where X_i is the vector of measurements for the i^{th} object. The most important special case occurs when (X_1, X_2, \dots, X_n) are independent and identically distributed. In this case, we have a **random sample** of size n from the common distribution.

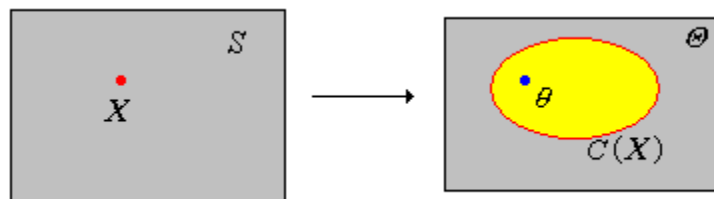
Suppose also that the distribution of X depends on a parameter θ taking values in a parameter space Θ . The parameter θ may also be vector-valued, in which case $\Theta \subseteq \mathbb{R}^k$ for some $k \in \mathbb{N}_+$ and $\theta = (\theta_1, \theta_2, \dots, \theta_k)$.

Confidence Sets

A **confidence set** is a subset $C(X)$ of the parameter space Θ that depends only on the data variable X , and no unknown parameters. Thus, in a sense, a confidence set is a *set-valued statistic*. A confidence set is an estimate of θ in the sense that we hope that $\theta \in C(X)$ with high probability. In particular, the **confidence level** is the smallest probability that $\theta \in C(X)$:

$$\min \{ \mathbb{P}_\theta(\theta \in C(X)) : \theta \in \Theta \}$$

Usually, we try to construct a confidence set for θ with a prescribed confidence level $1 - \alpha$ where $0 < \alpha < 1$. Typical confidence levels are 0.9, 0.95, and 0.99. Sometimes the best we can do is to construct a confidence set whose confidence level is *at least* $1 - \alpha$ this is called a **conservative** $1 - \alpha$ confidence set for θ .



Note that when we run the experiment and observe the data x , the *computed* confidence set is $C(x)$. The true value of the parameter θ is either in this set, or is not, and we will usually never know. However, by the **law**

of large numbers, if we were to repeat the confidence experiment over and over, the proportion of sets that contain θ would converge to $\mathbb{P}_\theta(\theta \in C(\mathbf{X})) \geq 1 - \alpha$. This is the precise meaning of the term **confidence**.

Next, note that the quality of a confidence set, as an estimator of θ , is based on two factors: the confidence level and the “size” of the set. A good estimate has small size (and hence gives tight bounds on θ) and large confidence. However, for a given \mathbf{X} , there is usually a tradeoff between confidence level and size--increasing the confidence level comes only at the expense of increasing the size of the set, and decreasing the size of the set comes only at the expense of decreasing the confidence level. How we measure the “size” of the confidence set depends on the dimension of the parameter space and the nature of the confidence set. Moreover, the size of the set is usually random, although in some special cases it may be deterministic.

▣ 1. Suppose that $C_i(\mathbf{X})$ is a $1 - \alpha_i$ level confidence set for θ for $i \in \{1, 2, \dots, k\}$. Show that if $\alpha = \alpha_1 + \alpha_2 + \dots + \alpha_k < 1$ then $C_1(\mathbf{X}) \cap C_2(\mathbf{X}) \cap \dots \cap C_k(\mathbf{X})$ is a conservative $1 - \alpha$ level confidence set for θ . *Hint:* Use [Bonferroni's inequality](#).

In many cases, we are interested in estimating a real parameter $\lambda = \lambda(\theta)$ taking values in an interval parameter space (a, b) . In this context our confidence set frequently has the form

$$C(\mathbf{X}) = \{\theta \in \Theta : L(\mathbf{X}) < \lambda(\theta) < U(\mathbf{X})\}$$

where $L(\mathbf{X})$ and $U(\mathbf{X})$ are statistics. In this case $(L(\mathbf{X}), U(\mathbf{X}))$ is called a **confidence interval** for λ . If $L(\mathbf{X})$ and $U(\mathbf{X})$ are both random, then the confidence interval is often said to be **two-sided**. In the special case that $U(\mathbf{X}) = b$ and $L(\mathbf{X})$ is random, $L(\mathbf{X})$ is called a **lower confidence bound** for λ and the interval $(L(\mathbf{X}), b)$ is called an **upper confidence interval** for λ . In the special case that $L(\mathbf{X}) = a$ and $U(\mathbf{X})$ is random, $U(\mathbf{X})$ is called a **upper confidence bound** for λ and the interval $(a, U(\mathbf{X}))$ is called an **lower confidence interval** for λ .

▣ 2. Suppose that $L(\mathbf{X})$ is a $1 - \alpha$ level confidence lower bound for λ and that $U(\mathbf{X})$ is a $1 - \beta$ level confidence upper bound for λ . Show that if $\alpha + \beta < 1$ then $(L(\mathbf{X}), U(\mathbf{X}))$ is a conservative $1 - (\alpha + \beta)$ level confidence interval for λ . *Hint:* Use [Exercise 1](#)

Pivot Variables

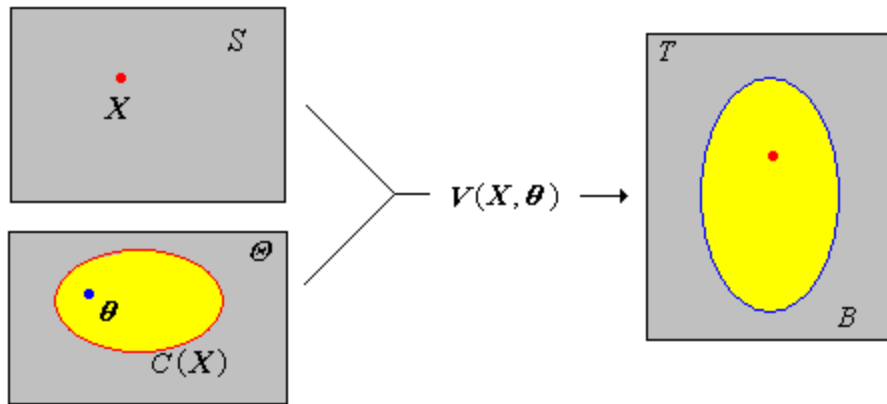
You might think that it should be very difficult to construct confidence sets for a parameter θ . However, in many important special cases, confidence sets can be constructed easily from certain random variables known as *pivot variables*.

Suppose that V is a function from $S \times \Theta$ into a set T . The random variable $V(\mathbf{X}, \theta)$ is a **pivot variable** for θ if its distribution does not depend on θ . Specifically, $\mathbb{P}_\theta(V(\mathbf{X}, \theta) \in B)$ is constant in $\theta \in \Theta$ for each $B \subseteq T$. If we know the distribution of the pivot variable, then for a given α , we can try to find $B \subseteq T$ (that does not depend on θ) such that

$$\mathbb{P}_\theta(V(\mathbf{X}, \theta) \in B) = 1 - \alpha$$

It then follows that a $1 - \alpha$ confidence set for the parameter is given by

$$C(X) = \{\theta \in \Theta : V(X, \theta) \in B\}$$

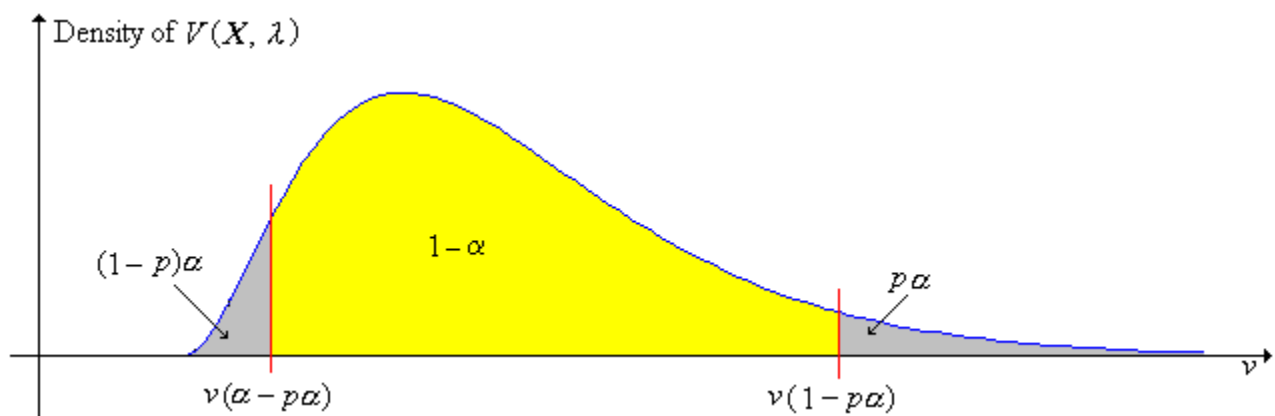


Suppose now that our pivot variable $V(X, \theta)$, is real-valued, which for simplicity, we will assume has a **continuous distribution**. For $p \in (0, 1)$, let $v(p)$ denote the **quantile** of order p for the pivot variable $V(X, \theta)$. By the very meaning of pivot variable, $v(p)$ does not depend on θ .

3. Show that for any $p \in (0, 1)$, a $1 - \alpha$ level confidence set for θ is

$$\{\theta \in \Theta : v(\alpha - p\alpha) < V(X, \theta) < v(1 - p\alpha)\}$$

The confidence set in [Exercise 3](#) corresponds to $(1 - p)\alpha$ in the left tail and $p\alpha$ in the right tail, in terms of the distribution of the pivot variable $V(X, \lambda)$. The special case $p = \frac{1}{2}$ is the **equal-tailed case**, the most common case.



4. Show that the confidence set in [Exercise 3](#) is decreasing in α and hence increasing in $1 - \alpha$ (in the sense of the subset relation) for fixed p .

Specializing further, suppose that $\theta = (\theta_1, \theta_1, \dots, \theta_n)$ is a vector of real parameters, and that we are interested in estimating one of the coordinates θ_i of θ ; the other coordinates are sometimes referred to as **nuisance parameters** in this context. It is often the case that the real-valued pivot variable $V(x, \theta)$ is a strictly decreasing function of θ_i for each $x \in S$ and for all values of the other coordinates of θ . In this setting, we can obtain a confidence set by inverting the pivot variable with respect to θ_i .

5. In the setting above, show that the $1 - \alpha$ confidence set for θ in [Exercise 3](#) can be written in the following form, where $\bar{\theta}_i$ is the parameter vector θ with θ_i deleted:

$$\{\theta \in \Theta : W(X, \bar{\theta}_i, v(1 - p\alpha)) < \theta_i < W(X, \bar{\theta}_i, v(\alpha - p\alpha))\}$$

In words, we apply the inverse transformation to obtain bounds on θ_i that depend on the data variable X , the other coordinates (nuisance parameters) of θ , and the quantiles of the pivot variable. If the other coordinates are known, then these bounds become statistics, and we have constructed a confidence *interval* for θ_i .

For the confidence set in [Exercise 3](#), we would naturally like to choose p that minimizes the size of the set in some sense. However this is often a difficult problem. The equal-tailed interval, corresponding to $p = \frac{1}{2}$, is the most commonly used case, and is sometimes (but not always) an optimal choice.

Pivot variables are far from unique; the challenge is to find a pivot quantity whose distribution is known and which gives tight bounds on the parameter.

6. Suppose that V is a pivot variable for θ . If g is a function defined on the range of V and g involves no unknown parameters, show that $U = g(V)$ is also a pivot variable for θ .

Location-Scale Families

In the case of [location-scale families](#) of distributions, we can easily find pivot variables. Suppose that Z is a real-valued random variable with a continuous distribution that has probability density function g , and no unknown parameters. Let $X = \mu + \sigma Z$ where $\mu \in \mathbb{R}$ and $\sigma \in (0, \infty)$ are parameters. Recall that the probability density function of X is given by

$$f_{\mu, \sigma}(x) = \frac{1}{\sigma} g\left(\frac{x - \mu}{\sigma}\right), \quad x \in \mathbb{R}$$

and the corresponding family of distributions is called the **location-scale family** associated with the distribution of Z ; μ is the **location parameter** and σ is the **scale parameter**. Generally, we are assuming that these parameters are unknown.

Now suppose that $\mathbf{X} = (X_1, X_2, \dots, X_n)$ is a random sample of size n from the distribution of X ; this is our observable outcome vector. For each i , let

$$Z_i = \frac{X_i - \mu}{\sigma}$$

7. Show that $\mathbf{Z} = (Z_1, Z_2, \dots, Z_n)$ is a random sample of size n from the distribution of Z .

In particular, note that \mathbf{Z} is a pivot variable for (μ, σ) , since \mathbf{Z} is a function of \mathbf{X} , μ , and σ , but the distribution of \mathbf{Z} does not depend on μ and σ . Hence, any function of \mathbf{Z} will also be a pivot variable for (μ, σ) , (if the function does not involve the parameters). Of course, some of these pivot variables will be much more useful

than others in estimating μ and σ . In the following exercises, we will explore two common and important pivot variables.

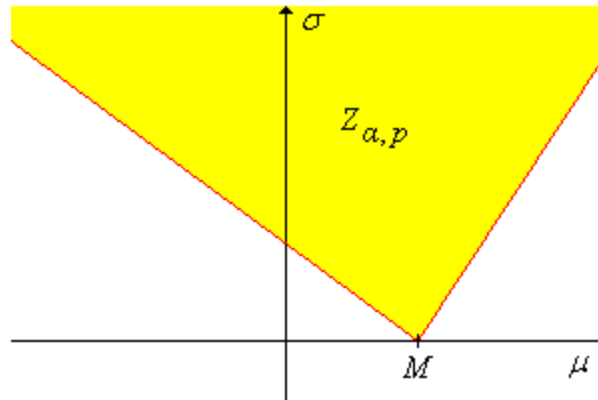
8. Let $M(X)$ and $M(Z)$ denote the **sample means** of X and Z , respectively. Show that $M(Z)$ is a pivot variable for (μ, σ) since

$$M(Z) = \frac{M(X) - \mu}{\sigma}$$

9. Let m denote the quantile function of the pivot variable $M(Z)$. Show that for any $p \in (0, 1)$, a $1 - \alpha$ confidence set for (μ, σ) is

$$Z_{\alpha,p}(X) = \{(\mu, \sigma) : M(X) - m(1 - p\alpha)\sigma < \mu < M(X) - m(\alpha - p\alpha)\sigma\}$$

10. Show that the confidence set in [Exercise 9](#) is a “cone” in the (μ, σ) parameter space, with vertex at $(M(X), 0)$ and boundary lines of slopes $-\frac{1}{m(1-p\alpha)}$ and $-\frac{1}{m(\alpha-p\alpha)}$, as shown in the graph below. (Note, however, that both slopes might be negative or both positive.)



The fact that the confidence set is unbounded is clearly not good, but is perhaps not surprising; we are estimating two real parameters with a single real-valued pivot variable. However, if σ is known, the confidence set defines a *confidence interval* for μ . Geometrically, the confidence interval simply corresponds to the horizontal cross section at σ .

11. In the confidence set in [Exercise 9](#), let $p \uparrow 1$ and $p \downarrow 0$, respectively, to show that $1 - \alpha$ confidence sets for (μ, σ) are

a. $Z_{\alpha,1}(X) = \{(\mu, \sigma) : M(X) - m(1 - \alpha)\sigma < \mu < \infty\}$

b. $Z_{\alpha,0}(X) = \{(\mu, \sigma) : -\infty < \mu < M(X) - m(\alpha)\sigma\}$

If σ is known, then [Exercise 11\(a\)](#) gives a $1 - \alpha$ confidence lower bound for μ and [Exercise 11\(b\)](#) gives a $1 - \alpha$ confidence upper bound for μ .

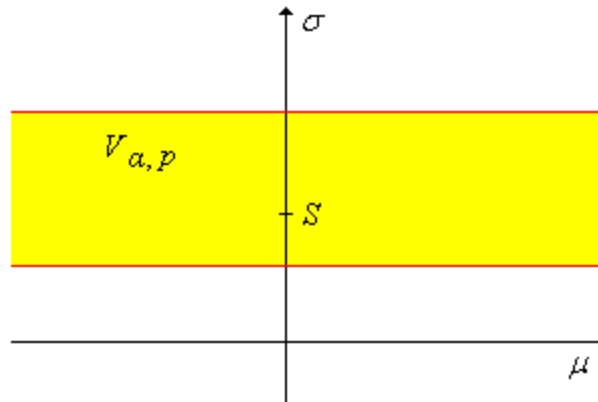
12. Let $S(X)$ and $S(Z)$ denote the **sample standard deviations** of X and Z , respectively. Show that $S(Z)$ is a pivot variable for (μ, σ) and a pivot variable for σ since

$$S(\mathbf{Z}) = \frac{S(\mathbf{X})}{\sigma}$$

13. Let s denote the quantile function of $S(\mathbf{Z})$. Use the pivot variable to show that for any $\alpha \in (0, 1)$ and any $p \in (0, 1)$, a $1 - \alpha$ confidence set for (μ, σ) is

$$V_{\alpha,p}(\mathbf{X}) = \left\{ (\mu, \sigma) : \frac{S(\mathbf{X})}{s(1-p\alpha)} < \sigma < \frac{S(\mathbf{X})}{s(\alpha-p\alpha)} \right\}$$

Note that the confidence set gives no information about μ since the random variable in [Exercise 13](#) is a pivot variable for σ alone. The confidence set can also be viewed as a bounded confidence *interval* for σ



14. In the confidence set in [Exercise 13](#), let $p \uparrow 1$ and $p \downarrow 0$, respectively, to show that $1 - \alpha$ confidence sets for (μ, σ) are

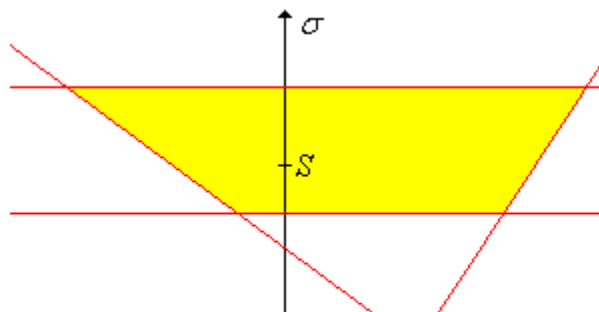
a. $V_{\alpha,1}(\mathbf{X}) = \left\{ (\mu, \sigma) : \frac{S(\mathbf{X})}{s(1-\alpha)} < \sigma < \infty \right\}$

b. $V_{\alpha,0}(\mathbf{X}) = \left\{ (\mu, \sigma) : 0 < \sigma < \frac{S(\mathbf{X})}{s(\alpha)} \right\}$

The set in part (a) gives a $1 - \alpha$ confidence lower bound for σ and the set in part (b) gives a $1 - \alpha$ confidence upper bound for σ

We can intersect the confidence sets corresponding to the two pivot variables to produce conservative, bounded confidence sets.

15. Suppose that $(\alpha, \beta, p, q) \in (0, 1)^4$ with $\alpha + \beta < 1$. Use [Exercise 1](#) to show that $Z_{\alpha,p}(\mathbf{X}) \cap V_{\beta,q}(\mathbf{X})$ is a conservative $1 - (\alpha + \beta)$ confidence set for (μ, σ)





The most important location-scale family is the family of [normal distributions](#). The problem of [estimation in the normal model](#) is considered in the next section. In the remainder of this section, we will explore another important scale family.

The Exponential Distribution

Suppose $X = (X_1, X_2, \dots, X_n)$ is a random sample of size n from the [exponential distribution](#) with scale parameter $\sigma \in (0, \infty)$. Let

$$Y = \sum_{i=1}^n X_i$$

16. Show that $\frac{2}{\sigma} Y$ has the [chi-square distribution](#) with $2n$ degrees of freedom, and hence is a pivot variable for σ .

Note that the variable in [Exercise 16](#) is a multiple of the variable in [Exercise 8](#) (with $\mu = 0$). Thus, let g_k denote the probability density function and G_k the distribution function for the chi-square distribution with k degrees of freedom. In addition, for $p \in (0, 1)$, let $\chi_k^2(p)$ denote the [quantile](#) of order p for the distribution. That is, $\chi_k^2(p) = G_k^{-1}(p)$. For selected values of k and p , $\chi_k^2(p)$ can be obtained from the [table of the chi-square distribution](#), from the [quantile applet](#), or from most statistical software packages.

17. Show or recall that

- $\chi_k^2(p) \rightarrow 0$ as $p \downarrow 0$
- $\chi_k^2(p) \rightarrow \infty$ as $p \uparrow 1$
- $\left(\frac{d\chi_k^2}{dp}\right)(p) = \frac{1}{g_k(\chi_k^2(p))}$ (*Hint: use the inverse function theorem of calculus.*)

18. Show that for any $\alpha \in (0, 1)$ and any $p \in (0, 1)$, a $1 - \alpha$ confidence interval for σ is

$$\left(\frac{2Y}{\chi_{2n}^2(1-p\alpha)}, \frac{2Y}{\chi_{2n}^2(\alpha-p\alpha)} \right)$$

19. Show that

- $\frac{2Y}{\chi_{2n}^2(1-\alpha)}$ is a $1 - \alpha$ confidence lower bound for σ .
- $\frac{2Y}{\chi_{2n}^2(\alpha)}$ is a $1 - \alpha$ confidence lower bound for σ .

Of the two-sided confidence intervals in [Exercise 18](#), we would naturally prefer the one with the smallest

length, because this interval gives the most information about the parameter b . However, minimizing the length as a function of p is computationally difficult. The two-sided confidence interval that is typically used is the **equal tailed** interval obtained by letting $p = \frac{1}{2}$:

$$\left(\frac{2Y}{\chi_{2n}^2\left(1 - \frac{\alpha}{2}\right)}, \frac{2Y}{\chi_{2n}^2\left(\frac{\alpha}{2}\right)} \right)$$

20. Try to find the p that minimizes the length of the interval in [Exercise 18](#).

[Virtual Laboratories](#) > [8. Set Estimation](#) > **1** [2](#) [3](#) [4](#) [5](#)

[Contents](#) | [Applets](#) | [Data Sets](#) | [Biographies](#) | [External Resources](#) | [Key words](#) | [Feedback](#) | ©