Interval Estimation Under The Uniform Distribution U(a,b)

QIQING YU
Department of Mathematical Sciences
State University of New York
Binghamton, NY 13902
USA

Abstract: - In this short note, we consider interval estimation for the parameters under the uniform distribution $U(a, b)$. We study two approaches: (1) based on a Wald-type statistic, (2) based on a pivotal statistic. We show that the first approach in its common form is not valid and we propose a modified version of the first approach. It turns out it is equivalent to the confidence interval with the shortest length.

Key-Words: - Maximum likelihood estimator, Confidence Intervals, Pivotal Statistic, Wald-type Statistic

Received: April 20, 2021. Revised: January 7, 2022. Accepted: January 25, 2022. Published: February 24, 2022.

1 Introduction.

The uniform distribution $U(a, b)$ is a common distribution and has been studied extensively (see, for example, Kuipers and Niederreiter (2012), Stephens (2017) and Claessen. *et al.* (2015), among others. The maximum likelihood estimators (MLEs) of its parameters have explicit expressions. How to construct a confidence interval (CI) under $U(a, b)$ is a typical content in a basic statistics course. For example, in the textbook by Casella and Bergera (2002), it is explained that if the data are from $U(0, b)$ then the exact CI for b can be constructed using a pivotal statistic. If the random sample is from $U(a, b)$ when both a and b are parameters, then a, b and θ are parameters, where $\theta = b - a$. Under this assumption, we shall show that there does not exist an exact CI. We shall discuss how to construct approximate CIs.

2 Theory.

Let W_1, ..., W_n be i.i.d. from $W \sim U(a, b)$, with the cumulative distribution function (cdf) $F_W(\cdot)$. Let $(\hat{a}, \hat{b}, \hat{\theta})$ be the MLE of (a, b, θ), where $\hat{a} = W_{(1)} = \min_i W_i$, $\hat{b} = W_{(n)} = \max_i W_i$ and $\hat{\theta} = \hat{b} - \hat{a}$. Recall that $P(\hat{b} \le t) = P(W_{(n)} \le t) = P(W_i \le t,\ \forall\ i) = (F_W(t))^n$ and $P(W_{(1)} > t) = P(W_i > t,\ \forall\ i) = (S_W(t))^n$, where $S_W = 1 - F_W$. Thus the distribution of the MLE of (a, b, θ) is well understood. It is easy to verify that $(\hat{a}, \hat{b}, \hat{\theta})$ is consistent.

There are two possible approaches in constructing CI's for $\gamma \in \{a, b, \theta\}$: (1) base on the Wald-type statistic $\frac{\hat{\gamma}-\gamma}{\hat{\sigma}_{\hat{\gamma}}}$, *e.g.*, $[\hat{\gamma} - 1.96\hat{\sigma}_{\hat{\gamma}},, \hat{\gamma} + 1.96\hat{\sigma}_{\hat{\gamma}}]$, (2) based a pivotal statistic $T = \frac{W-a}{b-a}$, where $T \sim U(0, 1)$.

The first approach relies on the mean and variance. Recall the cdfs and density functions:
$F_{T_{(n)}}(t) = t^n$, $f_{T_{(n)}}(t) = nt^{n-1}$, $F_{T_{(1)}} = 1 - S_{T_{(1)}}$, $S_{T_{(1)}}(t) = (1 - t)^n, f_{T_{(1)}}(t) = n(1 - t)^{n-1}$, for $t \in [0, 1]$; and for $t, s \in (0, 1)$,

$$f_{T_{(1)},T_{(n)}}(t, s) = \frac{n! f_T(t)(F_T(s) - F_T(t))^{n-2} f_T(s)}{1!(n-2)!1!}.$$

Based on $W = \theta T + a$, it is easy to derive

$$\sigma^2_{\hat{b}} = \theta^2 \sigma^2_{T_{(n)}} = \sigma^2_{\hat{a}} = \frac{\theta^2 n}{(n+1)^2(n+2)}$$

$$and\ \ \sigma^2_{\hat{\theta}} = \theta^2 \sigma^2_{T_{(n)}-T_{(1)}} = \frac{2(n-1)\theta^2}{(n+2)(n+1)^2}. \quad (1)$$

The proofs are also given in Appendix.

3 The Main Results.

We shall consider constructing the CI for a, b or θ under the assumption that $W \sim U(a, b)$. For simplicity, we only discuss the case of a 95% CI (or $(1-\alpha)100\%$ CI's, with $\alpha = 0.05$). For general $(1 - \alpha)100\%$ CI's, just replace 0.05 by α.

3.1. CIs for b: First consider the pivotal method. Since $T = \frac{W-a}{b-a} \sim U(0, 1)$, T is a pivotal statistic. For $t \in [0, 0.05]$, $F_{T_{(n)}}(t) = t^n$, letting $(u, v) = (t^{1/n}, (0.95 + t)^{1/n})$ yields

$$0.95 = P(u \le T_{(n)} \le v) = P(u \le \tfrac{W_{(n)}-a}{b-a} \le v)$$
$$= P(\tfrac{1}{v} \le \tfrac{b-a}{W_{(n)}-a} \le \tfrac{1}{u})$$
$$= P(\tfrac{W_{(n)}-a}{v} + a \le b \le \tfrac{W_{(n)}-a}{u} + a).$$

If a is given, a 95% CI for b is

$$[\tfrac{W_{(n)}-a}{v} + a, \tfrac{W_{(n)}-a}{u} + a], \quad (2)$$

where $(u,v) = (t^{1/n},(0.95+t)^{1/n})$. There are 3 typical cases: $(u,v) = (0,0.95^{1/n})$, or $(0.025^{1/n},0.975^{1/n})$, or $(0.05^{1/n},1)$, with lengthes: $W_{(n)}(\frac{1}{0} - 0.95^{\frac{-1}{n}}, 0.025^{\frac{-1}{n}} - 0.975^{\frac{-1}{n}}, 0.05^{\frac{-1}{n}} - 1)$ $\approx W_{(n)}(\infty, 0.037, 0.030)$ if $n = 100$.
Thus the best choice among these three 95% CIs for b is $(W_{(n)}, \frac{W_{(n)}-a}{0.05^{1/n}} + a)$ if a is known. Actually, it is the shortest 95% CI, which is given by $(u,v) = (0.05^{1/n},1)$, as the length of the CI in Eq. (2) is $(W_{(n)} - a)[(0.95+t)^{-1/n} - t^{-1/n}]$ and

$((0.95+t)^{-1/n} - t^{-1/n})'_t$
$= \frac{-1}{n}(0.95+t)^{\frac{-1}{n}-1} - \frac{-1}{n}t^{\frac{-1}{n}-1} < 0$ for $t \in [0,0.05]$.
If a is unknown, estimating a by $W_{(1)}$ yields an approximate 95% CI

$$[W_{(n)}, \frac{W_{(n)} - W_{(1)}}{0.05^{1/n}} + W_{(1)}], \tag{3}$$

as $P(a < W_{(1)} \leq a+\delta) = P(T_{(1)} \leq \frac{\delta}{\theta})$
$= 1 - (1 - \frac{\delta}{\theta})^n \to 1 \ \forall\ \delta \in (0, \theta/2)$, and thus $P(W_{(n)} \leq b \leq \frac{W_{(n)}-W_{(1)}}{0.05^{1/n}} + W_{(1)}) \approx 0.95$ if n is large. The length of the CI is $(\hat{b} - \hat{a})(20^{1/n} - 1)$.

Wald-type statistic may lead to another possible 95% CI $\hat{b} \pm 1.96\hat{\sigma}_{\hat{b}}$, with its length $\approx 4\hat{\sigma}_{\hat{b}}$. By Eq. (1) and Eq. (3), if n is large then the ratio of these two lengths is

$\frac{(\hat{b}-a)(0.05^{-1/n}-1)}{4\hat{\sigma}_{\hat{b}}}$
$= \frac{(\hat{b}-a)(0.05^{-1/n}-1)}{4\hat{\theta}\sqrt{\frac{n}{n+2}}/(n+1)}$
$\approx \frac{0.05^{-1/n}-1}{4/n} < 0.8$,

thus $\hat{b} \pm 1.96\hat{\sigma}_{\hat{b}}$ is not as good as the CI in Eq. (3). Moreover, this approach is based on the belief that $P(\frac{W_{(n)}-b}{\hat{\sigma}_{W_{(n)}}} \leq t) \approx \Phi(t)\ \forall\ t$, where Φ is the cdf of $N(0,1)$. However, if $t > 0$ then

$0.95 \geq P(\frac{W_{(n)}-b}{\sigma_{W_{(n)}}} \leq t)$
$= P(W_{(n)} \leq t\sigma_{W_{(n)}} + b)$
$= (P(T \leq \frac{t\sigma_{W_{(n)}}+b-a}{b-a}))^n$
$= (P(T \leq \frac{t\sigma_{W_{(n)}}}{b-a} + 1))^n$
$= (1)^n$, as $T \sim U(0,1)$.

It leads to a contradiction: $0.95 \geq 1$. Thus $\hat{b} \pm 2\hat{\sigma}_{\hat{b}}$ is not a CI. But we can make use of the Wald-type statistic as follows. Choose $t < 0$ such that

$0.05 = P(\frac{W_{(n)}-b}{\sigma_{W_{(n)}}} \leq t)$
$= P(W_{(n)} \leq t\sigma_{W_{(n)}} + b)$
$= (P(W \leq t\sigma_{W_{(n)}} + b))^n$
$= (P(T \leq \frac{t\sigma_{W_{(n)}}+b-a}{b-a}))^n$
$= (\frac{t\frac{\theta\sqrt{\frac{n}{n+2}}}{n+1}}{\theta} + 1)^n \qquad (\approx (\frac{t}{n}+1)^n \approx e^t)$.

$0.95 = P(\frac{W_{(n)}-b}{\sigma_{W_{(n)}}} > t) \qquad (t \approx \ln 0.05 = -\ln 20)$
$= P(b < W_{(n)} - \sigma_{W_{(n)}}t)$
$= P(W_{(n)} < b < W_{(n)} - \sigma_{W_{(n)}}t)$
$\approx P(W_{(n)} < b < W_{(n)} + \sigma_{W_{(n)}}\ln 20)$.

Thus an approximate 95% CI for b is $(W_{(n)}, W_{(n)} + \hat{\sigma}_{W_{(n)}}\ln 20)$, with length $\approx \frac{\hat{b}-\hat{a}}{n}\ln 20 = (\hat{b} - \hat{a})\ln 20^{1/n}$, as $\sigma^2_{\hat{b}} = \frac{(b-a)^2 n}{(n+1)^2(n+2)}$ by Eq. (1). It is of interest to compare its length to the length of the CI in (3):

$(\hat{b} - \hat{a})(0.05^{-1/n} - 1)$
$= (\hat{b} - \hat{a})(20^{1/n} - 1)$
$\approx (\hat{b} - \hat{a})\ln 20^{1/n}$, as
$\ln 20^{1/n}$
$= \ln x$
$= \ln x - \ln 1$
$\approx (\ln x)'|_{x=1}(x-1)$
$= x - 1$ (with $x = 20^{1/n}$).

Thus these two approximate CI's have the same length asymptotically.

Remark. In general, an approximate $(1-\alpha)100\%$ CI for b is

$$\begin{cases} [W_{(n)}, W_{(n)} - \hat{\sigma}_{W_{(n)}}\ln\alpha] & \text{Wald-type method} \\ [W_{(n)}, \frac{\hat{\theta}}{\alpha^{1/n}} + W_{(1)}] & \text{pivotal method.} \end{cases}$$

3.2. CI for a**:** $\frac{W_{(1)}-a}{\theta} = T_{(1)}$ is a pivatol statistic and $P(T_{(1)} > t) = (1-t)^n$ if $t \in [0,1]$. Let (u,v) satisfy
$0.95 = P(u \leq \frac{W_{(1)}-a}{\theta} \leq v) \quad (= (1-u)^n - (1-v)^n)$
$0.95 = P(W_{(1)} - u\theta \geq a \geq W_{(1)} - v\theta)$, then it leads to an approximate 95% CI for a, *e.g,* let
$((1-u)^n, (1-v)^n)$
$= (0.95, 0), (0.975, 0.025), (1, 0.05)$, then $(u,v) = (1 - 0.95^{\frac{1}{n}}, 1)$ or $(1 - 0.975^{\frac{1}{n}}, 1 - 0.025^{\frac{1}{n}})$, or $(0, 1 - 0.05^{\frac{1}{n}})$.
Then their length
$= \theta(0.95^{\frac{1}{n}}, 0.975^{\frac{1}{n}} - 0.025^{\frac{1}{n}}, 1 - 0.05^{\frac{1}{n}})$. It can be shown that the shortest 95% CI for a is $[\hat{a} - v\theta, \hat{a}]$ if θ is given, where $v = 1 - 0.05^{1/n}$; otherwise, an approximate 95% CI is $[W_{(1)} - v\hat{\theta}, W_{(1)}]$.

Moreover, a 95% CI based on Wald-type statistic $\frac{\hat{a}-a}{\hat{\sigma}_{\hat{a}}}$ is $[\hat{a} - t\hat{\sigma}_{\hat{a}}, \hat{a}]$, where $t \approx n(1 - 0.05^{1/n})$ and $\hat{\sigma}_{\hat{\theta}} \approx \hat{\theta}/n$. The reason is as follows.

$0.95 = P(\frac{W_{(1)}-a}{\sigma_{W_{(1)}}} \leq t)$
$= P(W_{(1)} - t\sigma_{W_{(1)}} \leq a)$
$= P(W_{(1)} - t\sigma_{W_{(1)}} \leq a \leq W_{(1)})$
$= P(W_{(1)} \leq t\sigma_{W_{(1)}} + a)$
$= 1 - P(W_{(1)} > t\sigma_{W_{(1)}} + a)$
$= 1 - P(T_{(1)} > \frac{t\sigma_{W_{(1)}}+a-a}{\theta})$
$= 1 - P(T_{(1)} > \frac{t\sigma_{W_{(1)}}}{\theta})$
$= 1 - (1 - \frac{t\sigma_{W_{(1)}}}{\theta})^n$

$= 1-(1-\frac{t\theta\sqrt{\frac{n}{n+2}}}{\theta(n+1)})^n.$ =>

$0.05^{1/n} = 1 - t\frac{\sqrt{\frac{n}{n+2}}}{(n+1)}$ =>

$t = (1-0.05^{1/n})\Big/\frac{\sqrt{\frac{n}{n+2}}}{(n+1)} \approx n(1-0.05^{1/n}).$

Moreover, $\sigma_{W_{(1)}} = \frac{\theta\sqrt{\frac{n}{n+2}}}{(n+1)} \approx \theta/n.$

Remark. Both approaches lead to approximately the same CI, as expected.

3.3. CI for θ: Let $0.95 = 1-P(T_{(n)} \le v) = 1-v^n$, *i.e.* $v = 0.05^{1/n}$. Then

$P(T_{(n)} = \frac{W_{(n)}-a}{\theta} > v) = P(\frac{W_{(n)}-a}{v} > \theta)$

$= P(\frac{W_{(n)}-a}{v} > \theta > \hat{b}-\hat{a}).$

Hence, an approximate 95% CI for θ is $(\hat\theta, \frac{\hat\theta}{0.05^{1/n}}]$, where $\hat\theta = W_{(n)} - W_{(1)}$, Or in general, $(\hat\theta, \frac{\hat\theta}{\alpha^{1/n}}]$.

On the other hand, in order to study Wald-type approach, we need to find the distribution of $\hat\theta$ $(= W_{(n)} - W_{(1)})$. Let f be the density of $(T_{(n)}, T_{(1)})$.

$G(t) \overset{def}{=} P(T_{(n)} - T_{(1)} \le t)$

$= \int\int \mathbf{1}(0 \le x-y \le t) f(x,y)dxdy$

$= \int_0^t\int_0^x f(x,y)dydx + \int_t^1\int_{x-t}^x f(x,y)dydx \quad (t \in [0,1]).$

$G'(t) = \int_0^t f(t,y)dy - \int_0^t f(t,y)dy + \int_t^1 f(x,x-t)dx$

$= \int_t^1 n(n-1)t^{n-2}dx = n(n-1)t^{n-2}(1-t)$ =>

$G(t) = \int_0^t n(n-1)x^{n-2}(1-x)dx$

$= n(n-1)[\frac{x^{n-1}}{n-1} - \frac{x^n}{n}]|_0^t$ =>

$$G(t) = nt^{n-1} - (n-1)t^n, \quad t \in [0,1]. \tag{4}$$

Let v be determined by $0.05 = P(\frac{\hat\theta-\theta}{\sigma_{\hat\theta}} \le v)$. Then

$0.05 = P(\frac{\hat\theta-\theta}{\sigma_{\hat\theta}} \le v) = P(\hat\theta - v\sigma_{\hat\theta} \le \theta)$

$= P(W_{(n)} - W_{(1)} \le v\sigma_{\hat\theta} + \theta)$

$= P(T_{(n)} - T_{(1)} \le \frac{v\sigma_{\hat\theta}+\theta}{\theta})$

$= P(T_{(n)} - T_{(1)} \le \frac{v\frac{\theta\sqrt{\frac{2(n-1)}{n+2}}}{n+1}+\theta}{\theta})$

$= P(T_{(n)} - T_{(1)} \le \frac{v\sqrt{\frac{2(n-1)}{n+2}}}{n+1} + 1)$

$= G(v\frac{\sqrt{\frac{2(n-1)}{n+2}}}{n+1}) \quad (= 0.05);$

$0.95 = P(\hat\theta - v\sigma_{\hat\theta} > \theta)$

$= P(\hat\theta - v\sigma_{\hat\theta} > \theta \ge \hat\theta)$

$(= 1 - G(v\frac{\sqrt{\frac{2(n-1)}{n+2}}}{n+1})).$

Thus $[\hat\theta, \hat\theta - v\hat\sigma_{\hat\theta}]$ is an approximate 95% CI for θ, where v is specified by $G(v\frac{\sqrt{\frac{2(n-1)}{n+2}}}{n+1}) = 0.05$ and $G(t) = nt^{n-1} - (n-1)t^n$ by Eq. (4).

4 Summary.

The confidence intervals for the parameters under $U(a,b)$ are not of the typical form of $[\hat\psi - u, \hat\psi + u]$, but are of the form either $\hat\psi, \hat\psi+u]$, or $\hat\psi-u, \hat\psi]$, where $\psi \in \{a, b, \theta\}$.

Appendix

$E(T_{(1)}) = \frac{1}{n+1},$

$V(T_{(1)}) = \frac{n}{(n+1)^2(n+2)},$

$E(T_{(n)}) = \frac{n}{n+1}.$

$V(T_{(n)}) = \frac{n}{(n+1)^2(n+2)},$

$E(W_{(1)}) = \frac{b}{n+1} + a\frac{n}{n+1}.$

$E(W_{(n)}) = b\frac{n}{n+1} + \frac{a}{n+1}.$

$V(W_{(1)}) = V(W_{(n)}) = \frac{\theta^2 n}{(n+1)^2(n+2)},$

Let $Z = W_{(n)} - W_{(1)}$,

then $E(Z) = (b-a)(n-1)/(n+1).$

$\sigma_Z^2 = \frac{2n}{(n+1)^2(n+2)}\theta^2$

$-2[E(W_{(n)}W_{(1)}) - E(W_{(n)})E(W_{(1)})].$

$f_{W_{(i)},W_{(j)}}(x,y) =$

$\frac{n!(F_W(x))^{i-1}f_W(x)(F_W(y)-F_W(x))^{j-i-1}f_W(y)(S_W(y))^{n-j}}{(i-1)!1!(j-i-1)!1!(n-j)!},$

$x < y, i < j,$

$f_{W_{(1)},W_{(n)}}(x,y) = \frac{n!f_W(x)(F_W(y)-F_W(x))^{n-2}f_W(y)}{(n-2)!}$

$= n(n-1)\frac{(y-x)^{n-2}}{\theta^2}, a < x < y < b.$

$\sigma^2_{\hat\theta} = \sigma^2_{W_{(n)}-W_{(1)}}$

$= \frac{2n\theta^2}{(n+1)^2(n+2)} - \frac{2\theta^2}{(n+2)(n+1)^2} = \frac{2(n-1)\theta^2}{(n+2)(n+1)^2}.$

References:

[1] Casella, G. and Bergera, R. L. (2002). Statistical Inference (2nd ed.) Duxbury. U.S.

[2] Kuipers, L., & Niederreiter, H. (2012). Uniform distribution of sequences. Courier Corporation.

[3] Stephens, M. A. (2017). Tests for the uniform distribution. In Goodness-of-fit techniques (pp. 331-366). Routledge.

[4] Claessen, K., Duregard, J., & Palka, M. H. (2015). Generating constrained random data with uniform distribution. Journal of functional programming, 25.