カーネル型推定量を利用した推測について

中央大学理工学部　前園宜彦

1　カーネル型推定量

$X_1, X_2, ..., X_n$ を互いに独立で同じ分布にしたがう確率変数とし，分布関数を F_X，密度関数を f_X とする．この分布関数 F_X の推定量としては経験分布関数

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I(X_i \le x), \quad x \in \mathbf{R}, \tag{1}$$

がある．ここで $I(A)$ は定義関数である．経験分布関数は階段関数となり連続ではない．連続になるような推定量でよく利用されるのがカーネル型推定量である．

Parzen(1962) と Rosenblatt(1956) によるカーネル型密度関数は

$$\widehat{f}_X(x) = \frac{1}{nh}\sum_{i=1}^{n} k\left(\frac{x - X_i}{h}\right), \quad x \in \mathbf{R}, \tag{2}$$

で与えられる．ここで $k(\cdot)$ はカーネル関数と呼ばれ，$h > 0$ はスムーズさを制御するバンド幅と呼ばれるものである．k についてよく仮定されるのは原点について対称で，非負とし $\int_{-\infty}^{\infty} k(v)dv = 1$ の条件を満たすというものである．バンド幅は $n \to \infty$ のとき $h \to 0$ かつ $nh \to \infty$ とする．この密度関数推定量を積分すると，カーネル型分布関数推定量（Nadaraya (1964)）

$$\widehat{F}_X(x) = \frac{1}{n}\sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right), \quad x \in \mathbf{R}, \tag{3}$$

が得られる．ただし $K(v) = \int_{-\infty}^{v} k(u)du$ である．この推定量が一致性や漸近正規性を持つことは容易に示すことができる．$\widehat{F}_X(x)$ の漸近的性質については多くの研究がなされ，経験分布関数 $F_n(x)$ よりも良いことが示されている．適当な条件の下で $n \to \infty$ のとき

$$Bias[\widehat{F}_X(x)] = \frac{h^2}{2} f'_X(x) \int_{-\infty}^{\infty} v^2 k(v)dv + o(h^2), \tag{4}$$

$$Var[\widehat{F}_X(x)] = \frac{1}{n} F_X(x)[1 - F_X(x)] - \frac{2h}{n} r_1 f_X(x) + o\left(\frac{h}{n}\right) \tag{5}$$

となる．ただし $r_1 = \int_{-\infty}^{\infty} vk(v)K(v)dv$ で，多くの場合非負の値をとる．したがって平均積分二乗誤差は

$$\begin{aligned} MISE(\widehat{F}_X) &= \frac{h^4}{4}\left[\int_{-\infty}^{\infty} v^2 k(v)dv\right]^2 \int_{-\infty}^{\infty} [f'_X(x)]^2 dx \\ &\quad + \frac{1}{n}\int_{-\infty}^{\infty} F_X(x)[1 - F_X(x)]dx - \frac{2h}{n} r_1 + o\left(h^4 + \frac{h}{n}\right) \end{aligned}$$

で与えられる．

もしカーネル関数が $k(-u) = k(u)$ を満たせば，$r_1 > 0$ となることが示せる．多項式型カーネル関数を用いて最適なバンド幅のオーダーは $h = O(n^{-1/3})$ であることが示されている (Azzalini(1981)).

2　確率点推定

p-確率点は $Q(p) = F^{-1}(p)$ で定義される．確率点の推定は様々な分野で必要になっている．金融工学では VaR（バリュー・アット・リスク）に基づいて，自己資本を用意しなければならない BIS 規制がある．これは収益率の下側 1% のような稀な損失の確率点の推定となる．気象の分野では非常に稀な降水量に対して治水の政策を決定しないといけない．パラメトリックなモデルの下での確率点の推定では滑らかな推測結果を得ることが多い．他方ノンパラメトリックな推測は，経験分布関数の逆関数として求められる．しかし経験分布関数は離散で対応する推定量は滑らかではない．滑らかさを回復するものとしては，カーネル型分布関数推定量の逆関数を利用する方法がある．経験分布関数に基づく p-確率点 $Q(p)$ の推定量は

$$\widetilde{Q}(p) = F_n^{-1}(p) = \inf\{x; F_n(x) \geq p\}$$

で与えられる．$X_{1;n} \leq X_{2;n} \leq \cdots \leq X_{n;n}$ を順序統計量とすると，経験分布関数の性質から

$$\widetilde{Q}(p) \approx X_{[np]+1;n}$$

となる．ここで $[\cdot]$ はガウス記号である．一様分布 $U(0,1)$ からの無作為標本に基づく順序統計量を $U_{1;n} \leq U_{2;n} \leq \cdots \leq U_{n;n}$ とすると

$$X_{[np]+1;n} \equiv Q(U_{[np]+1;n})$$

の関係があり，一様分布の順序統計量の性質を利用して漸近的な期待値と分散を求めることができる．$Q(U_{[np]+1;n})$ はベータ分布 $Be([np]+1, n-[np])$ にしたがうことを使うと

$$E\left[\widetilde{Q}(p)\right] = Q(p) + O(n^{-1}), \quad V\left[\widetilde{Q}(p)\right] = \frac{p(1-p)}{nf^2(Q(p))} + O(n^{-2})$$

となる．よって漸近平均二乗誤差は

$$AMSE(\widetilde{Q}(p)) = \frac{p(1-p)}{nf^2(Q(p))} = \frac{\{Q'(p)\}^2 p(1-p)}{n}$$

で与えられる．

Falk(1984) や Maesono & Penev(2011) では，カーネル関数 $k(\cdot)$ を利用した滑らかな推定量

$$\widehat{Q}_{p,h} = \frac{1}{h}\int_0^1 F_n^{-1}(x) K\left(\frac{x-p}{h}\right) dx$$

の漸近分布について研究している．ここで h はバンド幅である．Maesono & Penev (2011) は 4 次オーダーカーネル

$$\int_{-\infty}^{\infty} u^j k(u) du = 0, \quad j = 1, 2, 3$$

を使って次の漸近表現を求めている．

[定理１] 4 次オーダーカーネル $k^{(4)}(x)$ がリプシッツ条件を満たし，サポートを $(-1,1)$ とする．また，$p \in (0,1)$ とし，その近傍で $Q^{(4)}(p)$ が存在し，$f(Q(p)) > 0$ が成り立つとする．さらにバンド幅を $h = n^{-1/4}(\log n)^{-1}$ とする．このとき

$$\widehat{Q}_{p,h} - Q^*(p) = \frac{1}{n}A_{1n} + \frac{1}{n^2 h}A_{2n} + \frac{\delta}{n} + o_P(n^{-1})$$

が成り立つ．ただし

$$\begin{aligned}
\delta &= Q'(p)(\frac{1}{2} - p) + \frac{1}{2}Q^{(2)}(p)p(1-p) \\
Q^*(p) &= \frac{1}{h}\int_0^1 F^{-1}(x)k\left(\frac{x-p}{h}\right)dx, \\
\hat{I}_x(Y_1) &= I(Y_1 \le p+h) - (p+h), \\
g_{1n}(Y_1) &= -\int_{-1}^1 Q'(p+hx)k(x)\hat{I}_x(Y_1)dx, \\
g_{2n}(Y_1, Y_2) &= -\int_{-1}^1 Q'(p+hx)k'(x)\hat{I}_x(Y_1)\hat{I}_x(Y_2)dx, \\
A_{1n} &= \sum_{i=1}^n g_{1n}(Y_i), \\
A_{2n} &= \sum_{C_{n,2}} g_{2n}(Y_i, Y_j)
\end{aligned}$$

である．この漸近表現を使うと $\widehat{Q}_{p,h}$ の漸近平均二乗誤差を求めることができる．

[定理２] （定理１）と同じ条件の下でカーネル型推定量 $\widehat{Q}_{p,h}$ の期待値と分散は

$$\begin{aligned}
E[\widehat{Q}_{p,h}] &= Q(p) + o(n^{-1}), \\
V[\widehat{Q}_{p,h}] &= \frac{\{Q'(p)\}^2 p(1-p)}{n} + \frac{2h}{n}Q'(p)^2\left(-\frac{1}{2} + \frac{1}{2}\int_{-1}^1 K^2(x)dx\right) \\
&\quad + o\left(\frac{h}{n}\right)
\end{aligned}$$

となる．よって漸近平均二乗誤差は

$$AMSE(\widehat{Q}_{p,h}) = \frac{\{Q'(p)\}^2 p(1-p)}{n} + \frac{2h}{n}Q'(p)^2\left(-\frac{1}{2} + \frac{1}{2}\int_{-1}^1 K^2(x)dx\right)$$

で与えられる．

したがって $a = -\frac{1}{2} + \frac{1}{2}\int_{-1}^1 K^2(x)dx$ の値が負であればカーネル型推定の方が良いということになる．$k(\cdot)$ が非負のカーネル関数であれば a の値は負である．今は 4 次カーネルを仮定しており，$k(x) < 0$ となる x もあるが，シミュレーションで利用したカーネル関数に対しては a の値は負である．

2.1 シミュレーション結果

$(n=40, p=0.1)$, $(n=50, p=0.1)$ と $(n=100, p=0.1)$ のもとでシミュレーションを行なった結果を表1から表3にまとめている．ただし仮定している母集団分布は，$N(0,1)$: 標準正規分布，$\exp(1)$: 指数分布および $\Gamma(2,2)$: ガンマ分布である．$\widehat{Q}(p)$ の MSE と $\widehat{Q}_{p,h}$ の MSE のシミュレーション結果である．
method1 は $\widehat{Q}(p)$ の MSE
method2 は 4 次カーネル $k_1(x)$ を用いた $\widehat{Q}_{p,h}$ の MSE
method3 は 4 次カーネル $k_2(x)$ を用いた $\widehat{Q}_{p,h}$ の MSE
ここで $k_1(x), k_2(x)$ は

$$\begin{aligned} k_1(x) &= \frac{315}{512}\left(11x^8 - 36x^6 + 42x^4 - 20x^2 + 3\right) I(|x|<1), \\ k_2(x) &= \frac{15}{32}\left(7x^4 - 10x^2 + 3\right) I(|x|<1) \end{aligned}$$

である．$k_2(x)$ はイパネクニコフ・カーネルを元に Jones & Signorini(1997) の方法で 4 次カーネル化した関数 である．また，バンド幅は $h = n^{-1/4}(\log n)^{-1}$ を使っている．

表 1: $n=40, p=0.1$

	$N(0,1)$	$\exp(1)$	$\Gamma(2,2)$
method1	0.06645949	0.00392274	0.10836327
method2	0.06144286	0.00282985	0.08742409
method3	**0.0611502**	**0.0028082**	**0.0865862**

表 2: $n=50, p=0.1$

	$N(0,1)$	$\exp(1)$	$\Gamma(2,2)$
method1	0.05806929	0.003141974	0.08249772
method2	0.05412691	0.002410130	0.06641359
method3	**0.05383951**	**0.002381212**	**0.06513428**

表 3: $n=100, p=0.1$

	$N(0,1)$	$\exp(1)$	$\Gamma(2,2)$
method1	0.02848164	0.001341913	0.03699864
method2	0.02714162	0.001145669	0.03230338
method3	**0.02693276**	**0.001134178**	**0.03198531**

3 順位検定の連続化への応用

順位に基づく検定では，統計量の分布が離散分布になるために，有意確率を計算し，その値が小さければ棄却するという方法で行うことが多い．これらの順位検定の良さにつ

いては漸近相対効率で比較するのが主流である．しかし Lehmann & D'abrera(2006) や Brown et al.(2001) でも指摘されているように，データの少しの変動で有意確率が大きく変動することがある．また以下に述べるように分布の刻みが小さい検定統計量に基づく有意確率の方が小さい値を取る傾向がある．ここでは一標本問題についてこの事実を確認し，問題を解決するためにカーネル法を用いた統計量の連続化について考察する．

$X_1, X_2, \ldots, X_n$ を互いに独立で同じ母集団分布 $F(x-\theta)$ にしたがう無作為標本とする．ここで対応する確率密度関数は $f(-x)=f(x)$ を満たす原点対称な分布とする．θ は未知母数で，帰無仮説 $H_0 : \theta = 0$ に対して対立仮説 $H_1 : \theta > 0$ の検定問題を考える．この問題に対して多くの順位検定統計量が提案されている．代表的なものは符号検定とウィルコクソンの符号付き順位検定である．これらの検定統計量は離散型の分布を持つために有意水準 α を設定して検定するやり方ではなく，有意確率を評価して，その値が十分に小さい時に帰無仮説を棄却する方法が良く利用される．この有意確率については Lehmann & D'abrera(2006) でも指摘されているように，検定統計量の取り得る値が細かいほど有意確率が小さくなる傾向がある．

$\psi(x) = 1\ (x \geq 0),\ = 0\ (x < 0)$ とおくとき，符号検定 S とウィルコクソンの符号付き順位検定 W は

$$
\begin{aligned}
S &= S(\boldsymbol{X}) = \sum_{i=1}^{n} \psi(X_i), \\
W &= W(\boldsymbol{X}) = \sum_{1 \leq i \leq j \leq n} \psi(X_i + X_j)
\end{aligned}
$$

と同値になる．ここで $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)^T$ である．観測値 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^T$ に対して $s = S(\boldsymbol{x})$ を計算し，もし有意確率 $P_0(S \geq s)$ が十分小さいと判断されるときは帰無仮説 H_0 を棄却することになる．同様に $w = W(\boldsymbol{x})$ に対して $P_0(W \geq w)$ が小さい時 H_0 を棄却する．ただし $P_0(\cdot)$ は帰無仮説の下での確率を表す．表 4 は有意確率が小さくなる裾の領域における有意確率の大小の割合を検証したものである．

$$
\Omega_{|x|} = \{\boldsymbol{x} \in \boldsymbol{R}^n \ \| \ |x_1| < |x_2| < \cdots < |x_n|\}
$$

とおき，標準正規分布の $(1-\alpha)$-点 $z_{1-\alpha}$ に対して，

$$
\Omega_\alpha = \left\{ \boldsymbol{x} \in \Omega_{|x|} \ \middle\| \ \frac{s - E_0(S)}{\sqrt{V_0(S)}} \geq z_{1-\alpha}, \quad \text{or} \quad \frac{w - E_0(W)}{\sqrt{V_0(W)}} \geq z_{1-\alpha} \right\}
$$

の裾の領域を考える．ただし $E_0(\cdot)$, $V_0(\cdot)$ は帰無仮説の下での期待値と分散を表す．ここで統計量の実現値は標本 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^T$ の成分の入れ替えについて不変であるから $|x_1| < |x_2| < \cdots < |x_n|$ の場合だけを考察して，2^n 個の符号の組み合わせ $\mathrm{sign}(x_i) = \pm 1 (i = 1, 2, \ldots, n)$ について数え上げれば良い．表 4 では W の有意確率が S の有意確率よりも小さい比率 W/S を表している．

表でもわかるように，もし小さい有意確率を得たいのであれば，ウィルコクソンの符号付き順位検定を使った方が有意確率が小さい場合が多く，なるべく大きい有意確率を得たいのであれば符号検定を使った方が良いというように，検定の恣意性が生じてくる．この理由は検定統計量の分布が離散分布であることに起因する．ここではカーネル法を利用した統計量の分布の連続化を提案しその性質を考察する．

表 4: S と W の有意確率の大小

	sample size	$n=10$	$n=20$	$n=30$
$z_{0.90}$	W/S	3.28	1.367	1.477
$z_{0.95}$	W/S	1.92	1.449	1.425
$z_{0.975}$	W/S	4.2	1.674	1.572

経験分布関数を $F_n(\cdot)$ とすると，定義関数 $I(\cdot)$ に対して

$$F_n(0) = \frac{1}{n}\sum_{i=1}^{n} I(X_i \le 0)$$

となり

$$S = n - nF_n(0-).$$

の関係式が成り立ち，符号検定との関連が分かる．経験分布関数のカーネル法に基づく平滑化は，カーネル型密度関数推定量を積分して

$$\widetilde{F}_n(0) = \frac{1}{n}\sum_{i=1}^{n} K\Big(\frac{0-X_i}{h_n}\Big)$$

が原点での推定量になる．ただし h はバンド幅で $h \to 0,\ nh \to \infty\ (n \to \infty)$ を満足する．したがって S の連続化として

$$\widetilde{S} = n - n\widetilde{F}_n(0) = n - \sum_{i=1}^{n} K\left(-\frac{X_i}{h}\right)$$

が考えられる．帰無仮説の H_0 の下で，$\widetilde{S}$ の漸近平均と漸近分散は母集団分布 $F(\cdot)$ に依存しないことが示される．

同様にウィルコクソンの符号付き順位検定統計量の連続化を得ることができる．簡単のためにマン・ホィットニー検定統計量

$$M = \sum_{i=1}^{n} \psi(X_i + Y_i)$$

を考察する．この統計量は $P\left(\frac{X_1+X_2}{2} > 0\right)$ の推定量と見なすことができるので，連続化として

$$\widetilde{W} = \frac{n(n+1)}{2} - \sum_{1\le i\le j\le n} K\left(-\frac{X_i+X_j}{2h}\right)$$

を提案する．この統計量の分布は帰無仮説の下でも母集団分布に依存することになるが，$\widetilde{S}$ と同様に漸近平均と漸近分散は $F(\cdot)$ に依存しない．この $\widetilde{S}$, $\widetilde{W}$ は次の定理で示されるように S, W と一次のオーダーの意味で同値になる．

[**定理３**] $f'(\cdot)$ が存在し $-\theta$ の周りで連続であると仮定する．バンド幅を $h = cn^{-d}(c > 0, \frac{1}{4} < d < \frac{1}{2})$ とする．ここで

$$0 < \lim_{n\to\infty} V_\theta\left[1 - K\left(-\frac{X_1}{h}\right)\right] < \infty,$$

$$0 < \lim_{n\to\infty} \mathrm{Cov}_\theta\left[1 - K\left(-\frac{X_1+X_2}{2h}\right), 1 - K\left(-\frac{X_1+X_3}{2h}\right)\right] < \infty$$

が成り立ち，$k(\cdot)$ は原点に対して対称なカーネルとする．このとき次が成り立つ．
(i) 標準化した符号検定統計量と標準化した平滑化符号検定は

$$\lim_{n\to\infty} E_\theta\left\{\frac{S - E_\theta(S)}{\sqrt{V_\theta(S)}} - \frac{\widetilde{S} - E_\theta(\widetilde{S})}{\sqrt{V_\theta(\widetilde{S})}}\right\}^2 = 0$$

を満たし，漸近的に同値である．
(ii) 標準化 W と $\widetilde{W}$ に対しては

$$\lim_{n\to\infty} E_\theta\left\{\frac{W - E_\theta(W)}{\sqrt{V_\theta(W)}} - \frac{\widetilde{W} - E_\theta(\widetilde{W})}{\sqrt{V_\theta(\widetilde{W})}}\right\}^2 = 0$$

が成り立ち，漸近的に同値である．

このことから $\widetilde{S}$ と $\widetilde{W}$ の漸近性規性が成り立ち，漸近相対効率も一致することが示せる．

$\widetilde{S}$ と $\widetilde{W}$ の分布は連続型の分布になっており，有意確率の偏りの問題はかなり改善される．これを検証するために S と W の有意確率の比較と同じように，統計量の分布の裾の部分

$$\widetilde{\Omega}_\alpha = \left\{\boldsymbol{x} \in \boldsymbol{R}^n \;\middle\|\; \frac{\widetilde{s}(\boldsymbol{x}) - E_0(\widetilde{S})}{\sqrt{V_0(\widetilde{S})}} \geq z_{1-\alpha}, \quad \text{or} \quad \frac{\widetilde{w}(\boldsymbol{x}) - E_0(\widetilde{W})}{\sqrt{V_0(\widetilde{W})}} \geq z_{1-\alpha}\right\}$$

での有意確率の比較を行ったのが表 5 である．$\widetilde{S}$ と $\widetilde{W}$ の分布は帰無仮説の下でも元の母集団分布に依存するので，ここでは標準化した統計量の正規分布での近似を利用してシミュレーションにより有意確率の比較を行った．表の数値は有意確率が小さくなった回数の比を表している．シミュレーションの繰り返し数は 100,000 回である．表 4 と比較すると，連続化したことにより，有意確率が一方的に小さい傾向を持つという問題点はかなり解消されていることが分かる．

4　境界バイアス縮小す推定量に基づくノンパラメトリック検定

境界バイアスを縮小するカーネル型推定量に基づくノンパラメトリック検定を考察する．元の分布のサポートが有界ない場合のカーネル型推定量のバイアスの縮小する下記の分布

表 5: $\widetilde{S}$ と $\widetilde{W}$ の有意確率の大小

	sample size	$n=10$	$n=20$	$n=30$
$z_{0.90}$	$\widetilde{W}/\widetilde{S}$	1.284	1.182	1.148
$z_{0.95}$	$\widetilde{W}/\widetilde{S}$	1.166	1.199	1.111
$z_{0.975}$	$\widetilde{W}/\widetilde{S}$	1.437	1.240	1.113

関数推定量を Rizky & Maesono(2022) は提案している．Ω を X_i の分布のサポートとし，g を $\mathbf{R} \to \Omega$ の既知の全単射関数とする．このとき $g^{-1}(X_1), g^{-1}(X_2), \cdots, g^{-1}(X_n)$ に対してカーネル型推定量を構築するものである．以下の条件を仮定する．

C1. カーネル関数 $k(v)$ は非負で連続で $v=0$ に関して対称

C2. 積分 $\int_{-\infty}^{\infty} v^2 k(v) \mathrm{d}v$ が存在し，$\int_{-\infty}^{\infty} k(v) \mathrm{d}v = 1$

C3. バンド幅 $h > 0$ は $h \to 0$ 及び $nh \to \infty$ を満たす $(n \to \infty)$

C4. 単調増加関数 g は $\mathbf{R}$ から Ω　への全単射

C5. 密度関数 f_X と関数 g は 2 回微分可能

F を仮定された分布関数とするとき，次の検定問題を考える．

$$\text{帰無仮説 } H_0 : F_X = F \qquad \text{対立仮設 } H_1 : F_X \neq F$$

このような一般的な設定の下でよく利用される検定として

$$KS_n = \sup_{x \in \mathbf{R}} |F_n(x) - F(x)|. \tag{6}$$

を検定統計量とするコルモゴロフ・スミルノフ検定がある．検定統計量 KS_n の実現値より大きくなる確率を数表より求めて，その値が α 以下の時に帰無仮説 H_0 を棄却することになる．同様にクラーメル・フォンミーゼス検定統計量

$$CvM_n = n \int_{-\infty}^{\infty} [F_n(x) - F(x)]^2 \mathrm{d}F(x), \tag{7}$$

も利用される．

この検定統計量の自然な拡張として，カーネル型推定量を利用した

$$\widehat{KS} = \sup_{x \in \mathbf{R}} |\widehat{F}_X(x) - F(x)| \tag{8}$$

および

$$\widehat{CvM} = n \int_{-\infty}^{\infty} [\widehat{F}(x) - F(x)]^2 \mathrm{d}F(x) \tag{9}$$

が考えられる．この統計量に対して Omelka et al. (2009, AS) は帰無仮説下で，通常のコルモゴロフ・スミルノフ検定統計量及びクラーメル・フォンミーゼス統計量と同じ分布に収束することを示した．

多くの文献で議論されているように，カーネル型推定量は分布のサポートが有界の時に境界バイアスを持つ．このバイアスは上記の検定統計量のよさにも影響を与える．Rizky & Maesono(2022) は全単射関数 g を利用したバイアスを縮小するカーネル型推定量

$$\widetilde{F}_X(x) = \frac{1}{n}\sum_{i=1}^{n} W\left(\frac{g^{-1}(x) - g^{-1}(X_i)}{h}\right), \quad x \in \Omega \tag{10}$$

を提案している．この $\widetilde{F}$ に対して

$$\widetilde{KS} = \sup_{x\in\mathbf{R}} |\widetilde{F}_X(x) - F(x)|, \tag{11}$$

$$\widetilde{CvM} = n\int_{-\infty}^{\infty} [\widetilde{F}_X(x) - F(x)]^2 \mathrm{d}F(x), \tag{12}$$

[定理３]
帰無仮説 $H_0 : F_X = F$ の下で

$$|KS_n - \widetilde{KS}| \to_p 0. \tag{13}$$

が成り立つ．また

$$|CvM_n - \widetilde{CvM}| \to_p 0. \tag{14}$$

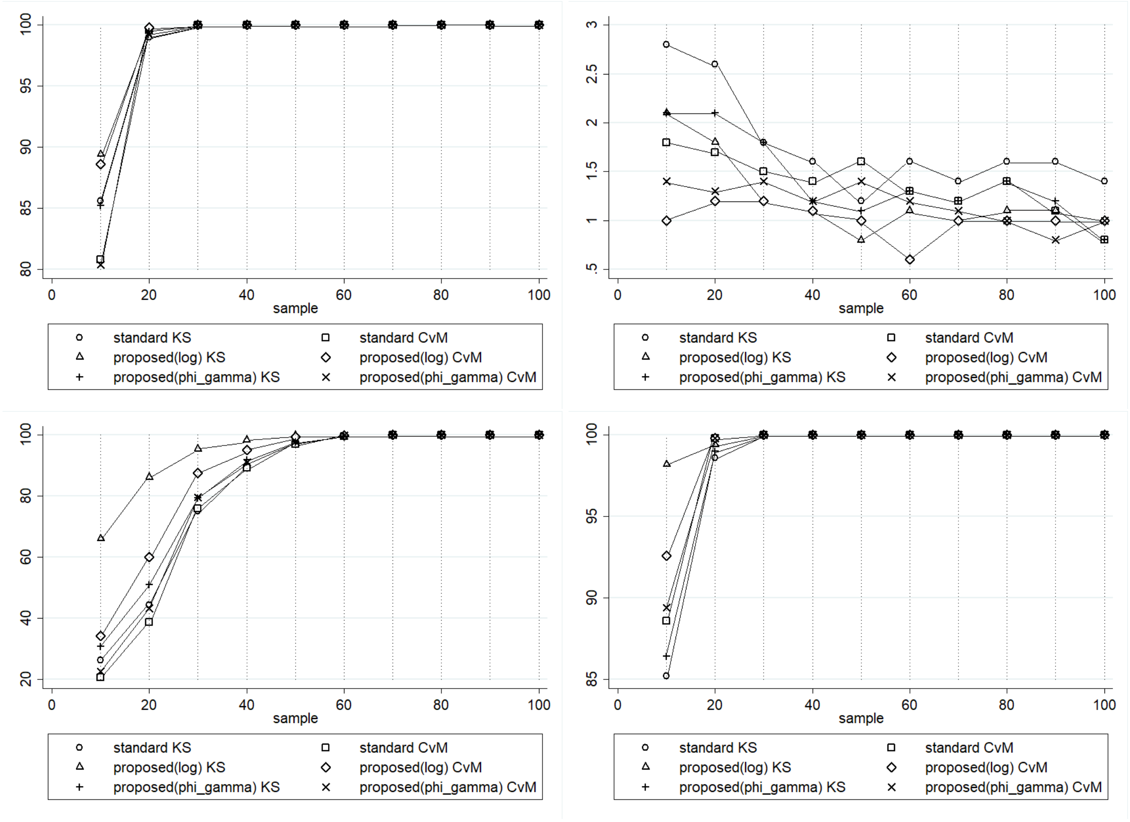
も成り立つ．

上記のことより，検定統計量 $\widetilde{KS}$ と $\widetilde{CvM}$ を利用した有意確率は，通常のコルモゴロフ・スミルノフ検定統計量およびクラーメル・フォンミーゼス検定統計量に対する数表を利用して検定できる．

[シミュレーションによる検出力の比較]
$n = 1, \sim, 100$ の標本数に対して，1000 回繰り返したときの検出力を図示している．有意水準は $\alpha = 0.01$ である．左上は帰無仮説が $H_0 : Gamma(2,2)\ vs.\ H_1 : Weibull$, 右上は 帰無仮説が $H_0 : Weibull(2,2)\ vs.\ H_1 : Weibull(2.2)$，左下は帰無仮説が $H_0 : logN(0,1)\ vs.\ H_1 : Weibull$ で右下は帰無仮説が $H_0 : absN(0,1)\ vs.\ H_1 : Weibull$ である．

参考文献

[1] Azzalini A. (1981), A note on the estimation of a distribution function and quantiles by a kernel method. *Biometrika*, **68**, 326-328.



[2] Brown, B., Hall, P. and Young, G. (2001). The smoothed median and the bootstrap. *Biometrika*, **88**, 519–534.

[3] Falk M. (1983). Relative efficiency and deficiency of kernel type estimators of smooth distribution functions. *Statist. Neerl.* **37**, 73-83.

[4] Jones M.C. and Signorini D.F. (1997). A Comparision of Higher-Order Bias Kernel Density Estimators. *Journal of American Statistical Association*, **92**,1063-1073.

[5] Lehmann E.L. and D'abrera H. (2006). *Nonparametrics: statistical methods based on ranks.* New York, Springer.

[6] Maesono Y. and Penev S. (2011). Edgeworth expansion for the kernel quantile estimator. *Annals of the Institute of Statistical Mathematics.* **63**, pp.617-644.

[7] Parzen E. (1962). On estimation of a probability density function and mode. *The Annals of Mathematical Statistics*, **32**, 1065-1076.

[8] Rizky R. F. and Maesono Y. (2022). Boundary-free estimators of the mean residual life function for data on general interval. submitted.

[9] Rosenblatt M. (1956). Remarks on some non-prametric estimates of a density function. *The Annals of Mathematical Statistics*, **27**, 832-837.