Statistics 583, Problem Set 4 Solutions

Wellner; 4/27/2011

1. Let $U_{m,n} \equiv T(\mathbb{F}_m, \mathbb{G}_n)$ where $T(F,G) = \int F dG = P(X \le Y)$ is the Mann-Whitney functional and $\mathbb{F}_m$ and $\mathbb{G}_n$ are the empirical df's of $X_1, \ldots, X_m$ i.i.d. with df F, $Y_1, \ldots, Y_n$ i.i.d. with df G where F and G are continuous.
(a) Show that

$$mnU_{m,n} + n(n+1)/2 = W_{m,n} \equiv \sum_{j=1}^{n} Q_j = \sum_{j=1}^{n} R_{m+j} \, .$$

(b) Show that $EU_{m,n} = P(X \le Y) = \int F dG$ and that

$$\begin{aligned} & Var(\sqrt{mn} U_{m,n}) \\ & = (n-1) \int (1-G)^2 dF + (m-1) \int F^2 dG - (N-1) \left(\int F dG \right)^2 + \int F dG \\ & = (n-1) Var[1 - G(X)] + (m-1) Var[F(Y)] + \int F dG \left(1 - \int F dG \right) . \end{aligned}$$

(c) Show that if $\lambda_N \equiv m/N \to \lambda \in [0,1]$, then, for independent standard Brownian bridge processes $\mathbb{U}$ and $\mathbb{V}$ it follows that

$$\begin{aligned} Var(\sqrt{mn/N} U_{m,n}) &= \frac{n-1}{N} Var(1 - G(X)) + \frac{m-1}{N} Var(F(Y)) + N^{-1} \int F dG \left(1 - \int F dG \right) \\ &\to (1-\lambda) Var(1 - G(X)) + \lambda Var(F(Y)) \\ &= (1-\lambda) \int \int (F(x) \wedge F(y) - F(x)F(y)) dG(x) dG(y) \\ &\quad + \lambda \int \int (G(x) \wedge G(y) - G(x)G(y)) dF(x) dF(y) \\ &= (1-\lambda) Var \left(\int \mathbb{U}(F) dG \right) + \lambda Var \left(\int \mathbb{V}(G) dF \right) \end{aligned}$$

as discussed in class on April 15. [Hint: the variance and covariance formulas in Chapter 1, Section 4, might be useful.]
(d) When $F = G$ use the results of A and B to compute $E_{(F,F)} W_{m,n}$ and $Var_{(F,F)}(W_{m,n})$. (This should agree with calculations for the Wilcoxon rank sum form of the statistic under the null hypothesis via finite sampling calculations.)

Solution: (a) Using empirical distribution function notation, $N\mathbb{H}_N = m\mathbb{F}_m + n\mathbb{G}_n$, so

$$\begin{aligned} mnU_{m,n} &= \int m\mathbb{F}_m d(n\mathbb{G}_n) = \int N\mathbb{H}_N d(n\mathbb{G}_n) - \int n\mathbb{G}_n d(n\mathbb{G}_n) \\ &= \sum_{j=1}^{n} N\mathbb{H}_N(Y_j) - \sum_{j=1}^{n} n\mathbb{G}_n(Y_{(j)}) \\ &= \sum_{j=1}^{n} R_{m+j} - \sum_{j=1}^{n} j \\ &= \sum_{j=1}^{n} R_{m+j} - n(n+1)/2\,. \end{aligned}$$

(b) The expectation is easy:

$$E(U_{m,n}) = \frac{1}{mn}\sum_{j=1}^{m}\sum_{i=1}^{n} P(X_i \le Y_j) = P(X_1 \le Y_1) = \int F dG\,.$$

For the variance, we first calculate

$$\begin{aligned} E[mnU_{m,n}]^2 &= \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{m}\sum_{l=1}^{n} E1_{[X_i\le Y_j, X_k \le Y_l]} \\ &= \sum_{i=1}^{m}\sum_{j=1}^{n} E1_{[X_i \le Y_j]} + \sum_{i\ne k}\sum_{j=1}^{n} P(X_i \le Y_j, X_k \le Y_j) \\ &\qquad + \sum_{i=1}^{m}\sum_{j\ne l} P(X_i \le Y_j, X_i \le Y_l) + \sum_{i\ne k}\sum_{j \ne l} P(X_i \le Y_j, X_k \le Y_l) \\ &= mnP(X_1 \le Y_1) + m(m-1)nP(X_1 \le Y_1, X_2 \le Y_1) \\ &\qquad + mn(n-1)P(X_1 \le Y_1, X_1 \le Y_2) \\ &\qquad + m(m-1)n(n-1)P(X_1 \le Y_1, X_2 \le Y_2)\,, \end{aligned}$$

$$P(X_1 \le Y_1) = \int F dG,$$

$$P(X_1 \le Y_1, X_2 \le Y_1) = EP(X_1 \le Y_1, X_2 \le Y_1 | Y_1) = \int F^2(x) dG(x)\,,$$

$$P(X_1 \le Y_1, X_1 \le Y_2) = EP(X_1 \le Y_1, X_1 \le Y_2 | X_1) = \int (1 - G(x-))^2 dF(x),$$

and

$$P(X_1 \le Y_1, X_2 \le Y_2) = P(X_1 \le Y_1)^2 = (\int F dG)^2\,.$$

It follows by algebra that

$$
\begin{aligned}
\mathrm{Var}(mnU_{m,n}) &= E(mnU_{m,n})^2 - \{E(mnU_{m,n})\}^2 \\
&= mn\int FdG + m(m-1)n\int F^2dG \\
&\quad + mn(n-1)\int(1-G(x-))^2dF(x) \\
&\quad + m(m-1)n(n-1)\{\int FdG\}^2 - (mn\int FdG)^2 \\
&= m(m-1)n\{\int F^2dG - (\int FdG)^2\} \\
&\quad + mn(n-1)\{\int(1-G(x-))^2dF(x) - (\int FdG)^2\} \\
&\quad - mn\int FdG(1-\int FdG)\,.
\end{aligned}
$$

By noting that

$$
\begin{aligned}
\int FdG = P(X \le Y) &= 1 - P(X > Y) \\
&= 1 - \int G(x-)dF(x) = \int(1-G(x-))dF(x)\,,
\end{aligned}
$$

this yields the claimed variance formula (to within a left limit):

$$
\begin{aligned}
\mathrm{Var}(\sqrt{mn}U_{m,n}) &= (m-1)Var(F(Y)) + (n-1)Var(1-G(X-)) \\
&\quad + \int FdG\left(1-\int FdG\right)\,.
\end{aligned}
$$

(c) Dividing both sides in the last display by N yields

$$
\begin{aligned}
Var(\sqrt{mn/N}U_{m,n}) &= \frac{n-1}{N}Var(1-G(X)) + \frac{m-1}{N}Var(F(Y)) + N^{-1}\int FdG\left(1-\int FdG\right) \\
&\to (1-\lambda)Var(1-G(X)) + \lambda Var(F(Y)) \\
&= (1-\lambda)\int\int(F(x)\wedge F(y) - F(x)F(y))dG(x)dG(y) \\
&\quad + \lambda\int\int(G(x)\wedge G(y) - G(x)G(y))dF(x)dF(y) \\
&\quad \text{by using (1.4.16) - (1.4.17) of Chapter 1, page 19} \\
&\quad \text{with } X \sim F \text{ and } h = G \\
&\quad \text{and then with } Y \sim G \text{ and } h = F \\
&= (1-\lambda)Var\left(\int \mathbb{U}(F)dG\right) + \lambda Var\left(\int \mathbb{V}(G)dF\right)
\end{aligned}
$$

since

$$
\begin{aligned}
&E\left(\int \mathbb{U}(F(x))dG(x)\right) = \int E\mathbb{U}(F(x))dG(x) = \int 0\cdot dG(x) = 0, \quad \text{and} \\
&E\left(\int \mathbb{U}(F(x))dG(x)\right)^2 = E\left(\int \mathbb{U}(F(x))dG(x)\int \mathbb{U}(F(y))dG(y)\right) \\
&\quad = E\int\int \mathbb{U}(F(x))\mathbb{U}(F(y))dG(x)dG(y) \\
&\quad = \int\int E\{\mathbb{U}(F(x))\mathbb{U}(F(y))\}dG(x)dG(y) \\
&\quad = \int\int \{F(x)\wedge F(y) - F(x)\cdot F(y)\}dG(x)dG(y)
\end{aligned}
$$

and similarly for $Var\left(\int \mathbb{V}(G(x))dF(x)\right)$. Here the interchanges of integration in (1) and (1) are justified by Fubini's theorem since

$$
\begin{aligned}
E|\int \mathbb{U}(F(x))dG(x)|^2 &\le E\int \mathbb{U}(F(x))^2dG(x) = \int E\mathbb{U}(F(x))^2dG(x) \\
&= \int F(x)(1-F(x))dG(x) \quad \text{by Tonelli's theorem} \\
&\le (1/4)\int dG(x) < 1/4.
\end{aligned}
$$

(d) When $F = G$ continuous we find that

$$E(U_{mn}) = \int FdF = 1/2\,,$$

and, since now $Var[F(Y)] = Var[G(X)] = 1/12$,

$$
\begin{aligned}
\mathrm{Var}(\sqrt{mn}U_{m,n}) &= (m-1)\frac{1}{12} + (n-1)\frac{1}{12} + \frac{1}{4} \\
&= (N-2)\frac{1}{12} + \frac{1}{4} = (N+1)\frac{1}{12}\,.
\end{aligned}
$$

Hence from part (a) it follows that

$$E(\sum_{j=1}^{n} Q_j) = n(n+1)/2 + mnE(U_{m,n}) = n(N+1)/2$$

and

$$\mathrm{Var}(\sum_{j=1}^{n} Q_j) = mn\mathrm{Var}(\sqrt{mn}U_{m,n}) = mn(N+1)\frac{1}{12}$$

both of which agree with finite sampling calculations (drawing a sample of size n balls from an urn without replacement where the numbers on the N balls in the urn are $\{1, 2, \ldots, N\}$).

2. (See also van der Vaart (1998), page 303, problem 3.) For distribution functions F on R^+ and $t_0 > 0$, consider the functional $T(F) = \Lambda(t_0) \equiv \int_0^{t_0} \frac{1}{1-F_-} dF$, the *cumulative hazard function* corresponding to F at t_0.
(a) Find the influence function of $T(F)$.
(b) What does this mean about asymptotic normality of the natural estimator $T(\mathbb{F}_n)$ of $T(F)$?
(c) Can you prove asymptotic normality of $T(\mathbb{F}_n)$ directly?

Solution: (a) To find the influence function of $T(F)$, let $F_t = (1-t)F + t\delta_x$. The distribution function corresponding to $G \equiv \delta_x$ is $1_{(-\infty,y]}(x)$, $y \in R$, so the left limit is $G_-(y) = 1_{[x<y]}$, and the corresponding "at risk" function $1 - G_-(y) = 1_{[x \geq y]} = 1_{[y,\infty)}(x)$. We need to compute

$$
\begin{aligned}
&\lim_{t\to 0} \frac{T(F_t) - T(F)}{t} = \frac{d}{dt} T(F_t)|_{t=0} \equiv IC(x; T, F) \equiv \psi_F(x) \\
&= \frac{d}{dt}\{\int_0^{t_0} \frac{1}{1-(F_t)_-} dF_t\}|_{t=0} \\
&= \int_0^{t_0} \frac{1}{1-F_-} d(\delta_x - F) + \int_0^{t_0} \frac{(\delta_x - F_-)}{(1-F_-)^2} dF \\
&= \frac{1_{[0,t_0]}(x)}{1-F_-(x)} - \int_0^{t_0} \frac{1}{1-F_-} dF + \int_0^{t_0} \frac{1}{1-F_-} dF - \int_0^{t_0} \frac{(1-\delta_{x-})}{(1-F_-)^2} dF \\
&= \frac{1_{[0,t_0]}(x)}{1-F_-(x)} - \int_0^{t_0} \frac{1_{[x\geq y]}}{(1-F_-(y))^2} dF(y) \\
&= \frac{1_{[0,t_0]}(x)}{1-F_-(x)} - \int_0^{x} \frac{1_{[0,t_0]}(y)}{1-F_-(y)} d\Lambda(y) \\
&= \begin{cases} \frac{1}{1-F_-(x)} - \int_0^x \frac{1}{(1-F_-)^2} dF & \text{if} \quad x \leq t_0 \\ -\int_0^{t_0} \frac{1}{(1-F_-)^2} dF & \text{if} \quad x > t_0\,. \end{cases}
\end{aligned}
$$

The next to last formula for the influence function of $\Lambda(t_0)$ is natural from a martingale perspective. When F is continuous $F_- = F$, and the influence function computed above reduces to:

$$
\begin{aligned}
IC(x; T, F) &= 1_{[x\leq t_0]} - \frac{F(t_0)}{1-F(t_0)} 1_{[x>t_0]} \\
&= \frac{1_{[x\leq t_0]} - F(t_0)}{1-F(t_0)}\,.
\end{aligned}
$$

Note that $E_F \psi_F(X) = 0$ and (in the case of a continuous d.f. F)

$$E_F \psi_F^2(X) = \frac{F(t_0)}{1 - F(t_0)} .$$

(b) Gateaux differentiability as established in (a) gives useful information about the form of the linear term, but does not yield a proof of asymptotic normality. To prove asymptotic normality we can proceed in several ways:
(i) Establish some stronger form of differentiability such as Fréchet differentiability or Hadamard differentiability with respect to some metric d_* compatible with the empirical distribution function (or empirical measure).
(ii) Recognize some special structure associated with the functional under consideration.
(iii) Proceed directly by showing that the remainder term

$$\begin{aligned} R_n &\equiv \sqrt{n}(T(\mathbb{F}_n) - T(F)) - \dot{T}(F; \sqrt{n}(\mathbb{F}_n - F)) \\ &= \sqrt{n}(T(\mathbb{F}_n) - T(F)) - \int \psi_F(x) d\{\sqrt{n}(\mathbb{F}_n - F)(x)\} \end{aligned}$$

satisfies $R_n = o_p(1)$. We will take route (iii) in (c) below. For (i) in the current case, see van der Vaart (1998), *Asymptotic Statistics*, example 20.15, page 301 (and his Lemma 20.10, page 298). With regard to (ii) in the present case, note that for $X_{(n)} < t_0$ we have

$$\begin{aligned} \sqrt{n}(T(\mathbb{F}_n) - T(F)) &= \int_0^{t_0} \frac{1}{1 - \mathbb{F}_n(t-)} d\left(\mathbb{F}_n(t) - \int_0^t (1 - \mathbb{F}_n(s-)) d\Lambda(s)\right) \\ &= \int_0^{t_0} \frac{1}{1 - \mathbb{F}_n(t-)} d\mathbb{M}_n(t) \end{aligned}$$

where $\mathbb{M}_n(t) \equiv \sqrt{n}(\mathbb{F}_n(t) - \int_0^t (1 - \mathbb{F}_n(s-)) d\Lambda(s))$ is a mean-zero martingale, and hence the integral in the last display is also a martingale (in t_0). Then asymptotic normality (even as a process in t_0) follows from martingale CLTs; see e.g. Shorack & W (1986), chapters 6 and 7 and especially (7.1.4) page 295 and Theorem 7.5.2 page 312.

(c) To prove asymptotic normality of $T(\mathbb{F}_n)$ (assuming that F satisfies $F(t_0) <$

1), write

$$\begin{aligned}
\sqrt{n}(T(\mathbb{F}_n) - T(F)) &= \sqrt{n}\left\{\int_0^{t_0} \frac{1}{1-\mathbb{F}_n(s-)} d\mathbb{F}_n(s) - \int_0^{t_0} \frac{1}{1-F(s-)} dF(s)\right\} \\
&= \int_0^{t_0} \frac{1}{1-\mathbb{F}_n(s-)} d[\sqrt{n}(\mathbb{F}_n(s) - F(s))] \\
&\quad + \int_0^{t_0} \sqrt{n}\left\{\frac{1}{1-\mathbb{F}_n(s-)} - \frac{1}{1-F(s-)}\right\} dF(s) \\
&= \frac{\sqrt{n}(\mathbb{F}_n(t_0) - F(t_0))}{1-\mathbb{F}_n(t_0-)} - \int_0^{t_0} \sqrt{n}(\mathbb{F}_n(s) - F(s))^2 \frac{1}{(1-\mathbb{F}_n(s-))^2} d\mathbb{F}_n(s) \\
&\quad + \int_0^{t_0} \frac{\sqrt{n}(\mathbb{F}_n(s) - F(s))}{(1-\mathbb{F}_n(s-))(1-F(s-))} dF(s) \\
&= \frac{\sqrt{n}(\mathbb{F}_n(t_0) - F(t_0))}{1-\mathbb{F}_n(t_0-)} + o_p(1) \\
&\to_d \frac{\mathbb{U}(F(t_0))}{1-F(t_0-)} \\
&\sim N(0, \frac{F(t_0)}{1-F(t_0)}) \qquad \text{if } F \text{ is continuous}
\end{aligned}$$

since the last two terms can be rewritten as

$$\int_0^{t_0} \frac{\sqrt{n}(\mathbb{F}_n(s) - F(s))}{1-\mathbb{F}_n(s-)} \left\{\frac{d\mathbb{F}_n(s)}{1-\mathbb{F}_n(s-)} - \frac{dF(s)}{1-F(s-)}\right\} = o_p(1)$$

by arguments similar to those we used to deal with the Mann-Whitney Wilcoxon statistic. Alternatively, martingale methods also work.

3. Let F be a distribution function on $\mathbb{R}^2$ with finite second moments, and let $\rho(F)$ be the correlation coefficient

$$\rho(F) = \frac{Cov_F(X,Y)}{\sqrt{Var_F(X) Var_F(Y)}}.$$

Assume that $|\rho(F)| < 1$.
(a) Give an example of a sequence of bivariate distributions $\{F_n\}$ satisfying $F_n \to_d F$, but $\rho(F_n) \to 1 \neq \rho(F)$.
(b) Find a collection $\mathcal{F}$ of distribution functions on $\mathbb{R}^2$ so that ρ is weakly continuous on $\mathcal{F}$.

Solution: (a) Without loss of generality we may suppose that F is a bivariate distribution function with zero means, $E_F(X) = E_F(Y) = 0$. Let $F_n = (1 -$

$n^{-1})F + n^{-1}\delta_{(a_n,b_n)}$ with $(a_n, b_n) \in \mathbb{R}^2$. Note that F_n has marginal distribution functions $F_{n,X} = (1-n^{-1})F_X + n^{-1}\delta_{a_n}$, $F_{n,Y} = (1-n^{-1})F_Y + n^{-1}\delta_{b_n}$ respectively where F_X and F_Y are the marginal df's of F. Thus we compute

$$\begin{aligned}
Cov_{F_n}(X,Y) &= E_{F_n}(XY) - E_{F_n}(X)E_{F_n}Y \\
&= (1-n^{-1})E(XY) + n^{-1}a_nb_n - ((1-n^{-1})E_FX + n^{-1}a_n)((1-n^{-1})E_FY + n^{-1}b_n) \\
&= (1-n^{-1})\{E_F(XY) - E_FX \cdot E_FY\} \\
&\quad + (1-n^{-1})E_FX \cdot E_FY - (1-n^{-1})^2 E_FX \cdot E_FY \\
&\quad - (1-n^{-1})E_FX \cdot n^{-1}b_n - (1-n^{-1})E_FY \cdot n^{-1}a_n - n^{-2}a_nb_n \\
&= (1-n^{-1})Cov_F(X,Y) + n^{-1}(1-n^{-1})a_nb_n \\
&\quad - n^{-1}(1-n^{-1})\{E_FX \cdot b_n + E_FY \cdot a_n\} + n^{-1}(1-n^{-1})E_FX \cdot E_FY \\
&= (1-n^{-1})Cov_F(X,Y) + n^{-1}(1-n^{-1})a_nb_n
\end{aligned}$$

since $E_FX = E_FY = 0$. Similarly,

$$\begin{aligned}
Var_{F_n}(X) &= E_{F_n}(X^2) - (E_{F_n}(X))^2 \\
&= (1-n^{-1})E_FX^2 + n^{-1}a_n^2 - ((1-n^{-1})E_FX + n^{-1}a_n)^2 \\
&= (1-n^{-1})\{E_F(X^2) - (E_FX)^2\} + n^{-1}(1-n^{-1})(E_FX)^2 \\
&\quad + n^{-1}(1-n^{-1})a_n^2 - 2n^{-1}(1-n^{-1})E_FX \cdot a_n \\
&= (1-n^{-1})Var_F(X) + n^{-1}(1-n^{-1})a_n^2 + n^{-1}(1-n^{-1})\{(E_FX)^2 - 2E_FX \cdot a_n\} \\
&= (1-n^{-1})Var_F(X) + n^{-1}(1-n^{-1})a_n^2, \qquad \text{and} \\
Var_{F_n}(Y) &= E_{F_n}(Y^2) - (E_{F_n}(Y))^2 = (1-n^{-1})Var_F(Y) + n^{-1}(1-n^{-1})b_n^2.
\end{aligned}$$

Choosing $a_n = b_n = n$ yields

$$\begin{aligned}
Cov_{F_n}(X,Y) &= n + o(n) = n(1+o(1)), \\
Var_{F_n}(X) &= n + o(n) = n(1+o(1)), \\
Var_{F_n}(Y) &= n + o(n) = n(1+o(1)).
\end{aligned}$$

Thus we find that

$$\rho(F_n) = \frac{Cov_{F_n}(X,Y)}{\sqrt{Var_{F_n}(X)Var_{F_n}(Y)}} = \frac{n(1+o(1))}{n(1+o(1))} \to 1$$

as $n \to \infty$. Thus ρ is weakly discontinuous at every F.

(b) Consider the following collection of distributions on R^2: for some $r > 2$ and $M < \infty$

$$\mathcal{F}_{r,M} \equiv \{F : \quad E_F|X|^r \le M, \; E_F|Y|^r \le M\}.$$

Then ρ is weakly-continuous on $\mathcal{F}_{r,M}$ at any F with $Var_F(X) > 0$ and $Var_F(Y) > 0$. Here is a proof: let $\{F_n\} \subset \mathcal{F}_{r,M}$ satisfy $F_n \to_d F$. Then with $(X_n, Y_n) \sim F_n$

and $(X, Y) \sim F$ we have $(X_n, Y_n) \to_d (X, Y)$, and by a Skorokhod construction there exist $(X_n^*, Y_n^*) =_d (X_n, Y_n)$ and $(X^*, Y^*) =_d (X, Y)$ defined on a common probability space and satisfying $(X_n^*, Y_n^*) \to_{a.s.} (X^*, Y^*)$. But because $\{F_n \subset \mathcal{F}_{r,M}$, X_n^2, Y_n^2, and $|X_n Y_n|$ are all uniformly integrable: since $r > 2$,

$$EX_n^2 1_{[X_n^2 \geq \lambda]} \quad \leq \quad \frac{1}{\lambda^{r-2}} E|X_n|^r \leq \frac{M}{\lambda^{r-2}}$$

so

$$\lim_{\lambda \to \infty} \limsup_{n \to \infty} EX_n^2 1_{[X_n^2 \geq \lambda]} \quad \leq \quad \lim_{\lambda \to \infty} \frac{M}{\lambda^{r-2}} = 0$$

and similarly for $\{Y_n^2\}$, so the uniform integrability of $|X_n Y_n|$ follows by Cauchy-Schwarz. The same holds true for the (X_n^*, Y_n^*) pairs since the uniform integrability only depends on the (marginal) distributions. Thus by Vitali's theorem it follows that

$$EX_n^s = EX_n^* s \to EX^* s = EX^s$$

and

$$EY_n^s = EY_n^* s \to EY^* s = EY^s$$

for $s = 1, 2$, while Vitali also yields

$$EX_n Y_n = EX_n^* Y_n^* \to EX^* Y^* = EXY\,.$$

Therefore

$$Var_{F_n}(X_n) \to Var_F(X), Var_{F_n}(Y_n) \to Var_F(Y), \tag{1}$$

and

$$Cov_{F_n}(X_n, Y_n) \to Cov_F(X, Y). \tag{2}$$

Since we have assumed that $Var_F(X) > 0$ and $Var_F(Y) > 0$, (1) and (2) yield

$$\rho(F_n) = \frac{Cov_{F_n}(X_n, Y_n)}{\sqrt{Var_{F_n}(X_n) \cdot Var_{F_n}(Y_n)}} \to \frac{Cov_F(X, Y)}{\sqrt{Var_F(X) \cdot Var_F(Y)}} = \rho(F);$$

i.e. ρ is continuous on $\mathcal{F}_{r,M}$ at any F with positive variances.

It is interesting to note that the hypothesis $\{F_n\} \subset \mathcal{F}_{r,M}$ *cannot be weakened to* $\{F_n\} \subset \mathcal{F}_{2,M}$ (and hence it can also not be weakened to the still larger class $\mathcal{F}_{2,\infty}$). Here is a counterexample. Let F be a d.f. on R^2 with $EX = 0 = EY$ and $EX^2 = 1 = EY^2$, and $\rho(F) < 1$ where $(X, Y) \sim F$. Let $M > 1$ be a big number, and consider the class

$$\mathcal{F}_{2,M} = \{F \text{ on } R^2 : \ E_F X^2 \leq M, E_F Y^2 \leq M\}\,.$$

Let $a_n, b_n > 0$; we will specify them in terms of M shortly. Consider the sequence of d.f.'s $\{F_n\} \subset \mathcal{F}_{2,M}$ defined by

$$F_n(x,y) = (1 - \frac{1}{n})F(x,y) + \frac{1}{2n}\delta_{(a_n,b_n)} + \frac{1}{2n}\delta_{(-a_n,-b_n)}.$$

Then for any bounded and continuous function $\psi : R^2 \to R$,

$$\begin{aligned} \int \psi dF_n &= (1 - \frac{1}{n}) \int \psi dF + \frac{1}{2n}\psi(a_n, b_n) + \frac{1}{2n}\psi(-a_n, -b_n) \\ &\to \int \psi dF, \end{aligned}$$

so $F_n \to_d F$. Furthermore, with $(X_n, Y_n) \sim F_n$,

$$EX_n = (1 - 1/n)EX = 0, EY_n = 0,$$

$$EX_n^2 = = (1 - 1/n)EX^2 + \frac{a_n^2}{n} = (1 - 1/n) + \frac{a_n^2}{n} = M$$

if $a_n^2 = n\{M - (1 - 1/n)\}$. Similarly,

$$EY_n^2 = (1 - 1/n) + \frac{b_n^2}{n} = M$$

if $b_n^2 = n\{M - (1 - 1/n)\}$. With these choices of a_n and b_n,

$$Cov(X_n, Y_n) = (1 - 1/n)Cov(X, Y) + \frac{a_n b_n}{n},$$

$$\begin{aligned} \rho(F_n) &= \frac{Cov(X_n, Y_n)}{Var(X_n)Var(Y_n)} \\ &= \frac{(1 - 1/n)Cov(X, Y) + M - (1 - 1/n)}{\sqrt{M^2}} \\ &\to \frac{\rho(F) + M - 1}{M} \neq \rho(F). \end{aligned}$$

Thus $\rho(F)$ is not continuous on $\mathcal{F}_{2,M}$.