1. A. Using empirical distribution function notation, $N\mathbb{H}_N = m\mathbb{F}_m + n\mathbb{G}_n$, so

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

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

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

where

$$\begin{aligned} P(X_1 \le Y_1) &= \int F\,dG\,, \\ P(X_1 \le Y_1\,, X_2 \le Y_{1)} &= EP(X_1 \le Y_1\,, X_2 \le Y_1 \,|\, Y_1) = \int F^2(x)\,dG(x)\,, \\ P(X_1 \le Y_1\,, X_1 \le Y_2) &= EP(X_1 \le Y_1\,, X_1 \le Y_2 \,|\, X_1) = \int (1 - G(x-))^2\,dF(x)\,, \end{aligned}$$

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

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

By noting that

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

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

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

C. When $F = G$ continuous we find that

$$E(U_{mn}) \;=\; \int F\, dF \;=\; 1/2\,,$$

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

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

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

$$Var(\sum_{j=1}^{n} Q_j) = mn\,Var(\sqrt{mn}\,U_{m,n})$$

$$= mn(N+1)\,\frac{1}{12}\,,$$

both of which agree with the finite sampling calculations of problem 4 of problem set 4.

2. A. If $\mathbf{F}_{3+\delta,M}$ is the collection of all distribution functions with $E_F|X|^{3+\delta} \le M$ for some $\delta > 0$ and $M < \infty$, then $|X|^k$ is uniformly integrable over $\mathbf{F}$ for $k = 0, 1, 2, 3$: for example,

$$E_F|X|^3 1_{\{|X|>K\}} \le E_F\{|X|^3 \frac{|X|^\delta}{K^\delta} 1_{\{|X|>K\}}\}$$

$$\le \frac{E_F\{|X|^{3+\delta}\}}{K^\delta} \le \frac{M}{K^\delta}$$

and hence

$$\lim_{K\to\infty} \sup_{F\in\mathbf{F}} E_F|X|^3 1_{\{|X|>K\}} = 0\,,$$

and this implies that if we have a sequence of distribution functions $\{F_n\} \subset \mathbf{F}_{3+\delta,M}$ with $F_n \to_d F$, then

$$E_{F_n}X^k \to E_F X^k\,, \qquad k = 1, 2, 3\,.$$

Hence it also holds for such a sequence $\{F_n\}$ that

$$\gamma_1(F_n) \to \gamma_1(F)\,.$$

For a complete arbitrary sequence $\{F_n\}$ with $F_n \to_d F$ there are no such guarantees, and in fact it is not hard to construct a sequence $\{F_n\}$ with

$$0 = \liminf_n \gamma_1(F_n) < \limsup_n \gamma_1(F_n) = \infty\,.$$

[Take $F_n = \Phi$, the standard normal df for n even, $F_n = (1-1/n)\Phi + (1/n)\delta_{n^{1/2}}$ for n odd.]

Similarly, if $\mathbf{F}_{4+\delta,M}$ is the collection of all distribution functions with $E_F|X|^{4+\delta} \le M$ for some $\delta > 0$ and $M < \infty$, then $|X|^k$ is uniformly integrable over $\mathbf{F}_{4+\delta,M}$ for $k = 0, 1, 2, 3, 4$, and this implies that if we have a sequence of distribution functions $\{F_n\} \subset \mathbf{F}_{4+\delta,M}$ with $F_n \to_d F$, then

$$E_{F_n}X^k \to E_F X^k\,, \qquad k = 1, 2, 3, 4\,.$$

Hence it also holds for such a sequence $\{F_n\}$ that

$$\gamma_2(F_n) \to \gamma_2(F)\,.$$

B. One way to proceed is to write the standard coefficient of skewness, $\gamma_1(F)$, in terms of F^{-1}:

$$\frac{E_F(X - \mu_F)^3}{\{Var_F(X)\}^{3/2}} = \frac{\int_0^1 [F^{-1}(t) - \int_0^1 F^{-1}(s)\,ds]^3\,dt}{\{\int_0^1 [F^{-1}(t) - \int_0^1 F^{-1}(s)\,ds]^2\,dt\}^{3/2}}.$$

This suggests defining a "trimmed skewness coefficient" γ_{1t} by

$$\gamma_{1t}(F) = \frac{\frac{1}{1-2\alpha}\int_\alpha^{1-\alpha} [F^{-1}(t) - (1-2\alpha)^{-1}\int_\alpha^{1-\alpha} F^{-1}(s)\,ds]^3\,dt}{\{\frac{1}{1-2\alpha}\int_\alpha^{1-\alpha} [F^{-1}(t) - (1-2\alpha)^{-1}\int_\alpha^{1-\alpha} F^{-1}(s)\,ds]^2\,dt\}^{3/2}}$$

Then, by the same argument we used in class to prove continuity of the trimmed mean, $\gamma_{1t}(F)$ is weakly continuous: if $F_n \rightarrow_d F$, then $\gamma_{1t}(F_n) \rightarrow \gamma_{1t}(F)$. Does $\gamma_{1t}(F)$ measure "skewness" of F? If F is symmetric about some point μ_s, then the trimmed mean $(1-2\alpha)^{-1}\int_\alpha^{1-\alpha} F^{-1}(t)\,dt = \mu_s$ and it is easy to see that $\gamma_{1t}(F) = 0$. Moreover, if F is asymmetric with $|F^{-1}(1/2-t) - \mu_s| \neq |F^{-1}(1/2+t) - \mu_s|$ for some t with $\alpha \le 1/2 - t < 1/2 + t \le 1-\alpha$, then $\gamma_{1t}(F) \neq 0$. But if the asymmetry of F occurs outside $[F^{-1}(\alpha), F^{-1}(1-\alpha)]$ then we can still have F asymmetric and $\gamma_{1t}(F) = 0$. Thus while γ_{1t} is not an ideal measure of skewness, it is weakly continuous: we have traded ease of interpretability for continuity.