Using Stein–Chen — an example: Consider a chessboard that measures 1 unit of length on each side, with N squares in total. You and I each pick m distinct squares, uniformly at random. For some subset A of the squares, with total area $|A|$ (and composed of $N|A|$ squares), let S_A be the number of squares in A that we *both* picked. If N is large and $m \approx \sqrt{\lambda N}$, then S_A is approximately Poisson with mean $\lambda|A|$.
Remark: Recall that a *Poisson point process* (PPP) with rate λ in a region U is a random collection of points with the property that if $N(A)$ is the number of points falling in the subset A of U, then for disjoint subsets $U_1, \dots U_n$, $N(U_1), \dots, N(U_n)$ are independent and Poisson distributed with means equal to $\lambda|U_1|, \dots, \lambda|U_n|$. It turns out that for a random collection of points to be a PPP, it suffices that for any subset A of U composed of a union of rectangles, the probability that A contains no points is $\exp(-\lambda|A|)$. This therefore (mostly) proves that the random set of locations we have both picked converges as $N \to \infty$ to a PPP on the chessboard. (more on this in a few weeks)
Proof:

For each k, let X_k be the indicator that we both picked the k^{th} square. First, we should check that

$$\mathbb{E}[S_A] = \sum_{k \in A} \mathbb{E}[X_k] = N|A| \left(\frac{m}{N}\right)^2 = \lambda|A|,$$

as promised. In particular, $p_k = \mathbb{E}[X_k] = (m/N)^2 = \lambda/N$.

We now define V_k, by defining a closely related set of picks: if we did not both pick square k, reassign randomly chosen picks to k as needed. To make this explicit, suppose that I pick squares $\mathbf{i} = \{i_1, i_2, \dots, i_m\}$, and you pick $\mathbf{j} = \{j_1, j_2, \dots, j_m\}$. We will define picks $\mathbf{i}'$ and $\mathbf{j}'$ by slightly adjusting $\mathbf{i}$ and $\mathbf{j}$. Let L and L' be iid numbers chosen uniformly from $\{1, 2, \dots, m\}$; then i_L and $j_{L'}$ will be the picks we rearrange if necessary. If $k \in \mathbf{i}$, then let $\mathbf{i} = \mathbf{i}'$. Otherwise, define $\mathbf{i}' = \mathbf{i} \setminus \{i_L\} \cup \{k\}$. Define $\mathbf{j}'$ in terms of $\mathbf{j}$ similarly, reallocating $j_{L'}$ if necessary. Then V_k is the size (cardinality) of the set $\mathbf{i}' \cap \mathbf{j}' \cap (A \setminus \{k\})$, namely, the number of resulting shared picks in A, excluding square k.

We can be slightly more explicit about checking that $\mathbf{i}'$ and $\mathbf{j}'$ have the correct distributions, namely, the distribution of $\mathbf{i}$ and $\mathbf{j}$ *given* that $k \in \mathbf{i} \cap \mathbf{j}$ (given that we both picked k). It suffices to check for just $\mathbf{i}$, and for $k = 1$. Since the distribution of $\mathbf{i}$ is invariant under permutations of $\{1, 2, \dots, N\}$, the distribution of $\mathbf{i}$ conditioned on the event $\{1 \in \mathbf{i}\}$ is invariant under permutations of $\{2, \dots, N\}$. This property also holds for $\mathbf{i}'$. Therefore, each have the same distribution, namely, that of $\{1\}$ along with a uniformly chosen collection of $m - 1$ numbers from $\{2, 3, \dots, N\}$. We also know therefore that V_k has the distribution of $S_A - 1$, conditioned on $X_k = 1$.

We now want to bound $\mathbb{E}[|S_A - V_k|]$. Note that S_A can differ from V_k in three ways: if $X_k = 1$; if my pick chosen to reallocate was matched to one of yours that lay in A; and if your pick chosen to reallocate was matched to one of mine that layin A. Then $\mathbb{E}[|S_A - V_k|] = \mathbb{E}[S_A - V_k]$ is no greater than the sum of the probabilities of these three events. More carefully, let U V, and W

be the respective indicators of these things,

$$U = X_k \tag{1}$$

$$V = \begin{cases} 1 & \text{if } i_L \in A \cap \mathbf{j}' \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

$$W = \begin{cases} 1 & \text{if } j_{L'} \in A \cap \mathbf{i}' \\ 0 & \text{otherwise} \end{cases}. \tag{3}$$

Then $|S_A - V_k| = S_A - V_k \leq U + V + W$, and so $\mathbb{E}[|S_A - V_k|] \leq \mathbb{E}[U] + \mathbb{E}[V] + \mathbb{E}[W]$. We know $\mathbb{E}[U] = (m/N)^2$, and since V depends on choosing one of S_A things out of a total of m, $\mathbb{E}[V|S_A] = \mathbb{E}[W|S_A] = S_A/m$, so

$$\mathbb{E}[V] = \sum_n \mathbb{P}\{S_A = n\}\mathbb{E}[V|S_A = n] = \sum_n \mathbb{P}\{S_A = n\}\frac{S_A}{m} = \frac{\mathbb{E}[S_A]}{m},$$

and so

$$\mathbb{E}[|S_A - V_k|] \leq \left(\frac{m}{N}\right)^2 + 2\frac{\mathbb{E}[S_A]}{m} = \frac{\lambda}{N} + 2\frac{\lambda|A|}{\sqrt{\lambda N}}$$

Therefore,

$$\begin{aligned} d_{\mathrm{TV}}(S_A, P) &\leq (1 \wedge (\lambda|A|)^{-1}) \sum_{k \in A} p_k \mathbb{E}[|S_A - V_k|] \\ &\leq (1 \wedge (\lambda|A|)^{-1}) N \frac{\lambda|A|}{N} \left(\frac{\lambda}{N} + 2\frac{\lambda|A|}{\sqrt{\lambda N}}\right) \\ &\leq 2(1 \wedge (\lambda|A|))|A|\sqrt{\frac{\lambda}{N}}. \end{aligned}$$