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 $\lambda$ 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 $U_1, \ldots, U_n$, $N(U_1), \ldots, N(U_n)$ are independent and Poisson distributed with means equal to $\lambda |U_1|, \ldots, \lambda |U_n|$. 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

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

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

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

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

We now want to bound $E[|S_A - V_k|]$. 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 lay in $A$. Then $E[|S_A - V_k|] = E[S_A - V_k]$ 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,

\[ U = X_k \quad (1) \]

\[ V = \begin{cases} 
1 & \text{if } i_L \in A \cap j' \\
0 & \text{otherwise} 
\end{cases} \quad (2) \]

\[ W = \begin{cases} 
1 & \text{if } j_L' \in A \cap j' \\
0 & \text{otherwise} 
\end{cases} \quad (3) \]

Then \(|S_A - V| = |S_A - V_k| \leq U + V + W\), and so \(E[|S_A - V_k|] \leq E[U] + E[V] + E[W]\).

We know \(E[U] = (m/N)^2\), and since \(V\) depends on choosing one of \(S_A\) things out of a total of \(m\), \(E[V|S_A] = E[W|S_A] = S_A/m\), so

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

and so

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

Therefore,

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