Probability approximations using Stein's method

Tutorial 2: Solutions

1. (a) We use the Stein equation for geometric approximation from Tutorial 1 to write, for $A \subseteq \mathbb{Z}^+$,

$$\begin{aligned}\mathbb{P}(W \in A) - \mathbb{P}(Z \in A) &= (1-p)\mathbb{E}[f(W+1)] - \mathbb{E}[f(W)] \\ &= (1-p)\mathbb{E}[f(W+1)] - (1-p)\mathbb{E}[f(W)|W>0] \\ &= (1-p)\mathbb{E}[f(W+1) - f(V+1)]\,,\end{aligned}$$

where f is as given in Q4 of Tutorial 1 and we use the fact that $f(0) = 0$.

Taking the supremum over $A \subseteq \mathbb{Z}^+$ then gives

$$d_{TV}(W,Z) \le (1-p) \sup_{A\subseteq\mathbb{Z}^+} \sup_{j,k\in\mathbb{Z}^+} |f(j) - f(k)|\mathbb{P}(W \neq V) \le \frac{1-p}{p}\mathbb{P}(W \neq V)\,,$$

where we have used the bound on f from part (b) of Q4 in Tutorial 1.

(b) Conditioning on W being positive is the same as conditioning on the number of summands N being at least 1. Since $(N|N>0)$ has the same distribution as $N+1$ (this is one way of writing the characterization at the heart of Stein's method for geometric approximation), we may construct V by writing

$$V + 1 = X + \sum_{i=1}^{N} X_i = X + W\,,$$

so that $V = W + X - 1$. Using the result of part (a) then gives

$$d_{TV}(W,Z) \le \frac{1-p}{p}\mathbb{P}(W \neq V) = \frac{1-p}{p}\mathbb{P}(X \neq 1)\,,$$

as required.

2. This question is asking us to prove Theorem 2.13 from the notes. We will use a very similar approach as in Theorem 2.12, with a modification to take account of the different

monotonicity condition.

$$\begin{aligned}
d_{TV}(W,Z) &\leq \frac{1}{\lambda}\sum_{i=1}^n p_i\mathbb{E}|U_i - V_i| = \frac{1}{\lambda}\sum_{i=1}^n p_i\mathbb{E}|U_i - X_i + X_i - V_i| \\
&\leq \frac{1}{\lambda}\sum_{i=1}^n p_i\left(\mathbb{E}|U_i - X_i - V_i| + \mathbb{E}|X_i|\right) \\
&= \frac{1}{\lambda}\sum_{i=1}^n p_i\left(\mathbb{E}[V_i + X_i - U_i] + \mathbb{E}[X_i]\right) \\
&= \frac{1}{\lambda}\sum_{i=1}^n p_i\left(\mathbb{E}[V_i] - \mathbb{E}[U_i] + 2\mathbb{E}[X_i]\right)\,.
\end{aligned}$$

We have already seen that $\sum_{i=1}^n p_i\mathbb{E}[U_i] = \lambda^2$ and $\sum_{i=1}^n \mathbb{E}[V_i] = \text{Var}(W) + \lambda^2 - \lambda$. Since

$$\sum_{i=1}^n p_i\mathbb{E}[X_i] = \sum_{i=1}^n p_i^2$$

the result follows on combining all these expressions.

3. In this question we are filling in the missing details from the birthday problem treated in Section 2.7 of the notes.

   (a) For each i, $\mathbb{P}(Y_i = 0) = (1 - 1/n)^m$, since in order for this event to occur we need all m people (independently) to be born on one of the $n-1$ days other than day i. Any given individual is born on one of these $n-1$ days with probability $1 - 1/n$. Linearity of expectation then gives that

   $$\mathbb{E}[W] = \sum_{i=1}^n \mathbb{E}[X_i] = \sum_{i=1}^n \mathbb{P}(Y_i = 0) = n\left(1 - \frac{1}{n}\right)^m\,,$$

   as required.

   (b) There are (at least) two ways of thinking about this calculation, depending on your taste. We can either

   - observe that the allocation of birthdays after the reassignment is the same probabilistically as if we had a one-stage assignment of birthdays, with m individuals assigned, independently and uniformly at random, to $n-1$ days (i.e, to one of the days other than day i); or
   - consider the original two-stage assignment and use a conditioning argument as follows: For each person, we can calculate the probability of them not being born on day $j \neq i$ after the stated reassignment. By considering separately the cases where (i) this person is initially assigned to neither day i nor day j, and (ii) this person is initially assigned to day i and then reassigned to a day

other than j, this probability is equal to

$$\left(1-\frac{2}{n}\right)+\frac{1}{n}\left(1-\frac{1}{n-1}\right)=1-\frac{1}{n-1}\,.$$

Since there are m people, each treated independently, the probability that no-one is born on day $j \neq i$ following the reassignment is therefore $(1-1/(n-1))^m$. There are $n-1$ such days, and as in part (a) we use the linearity of expectation.

Thankfully, these two approaches both give the same result!

4. Multiplying both sides of our Stein equation by the integrating factor $e^{-x^2/2}$ gives

$$e^{-x^2/2}\left(h(x)-\mathbb{E}h(Z)\right)=\frac{d}{dx}\left(e^{-x^2/2}f(x)\right)\,.$$

Integrating then gives

$$\begin{aligned} f(x) &= e^{x^2/2}\int_{-\infty}^{x}\left(h(y)-\mathbb{E}h(Z)\right)e^{-y^2/2}\,dy \\ &= -e^{x^2/2}\int_{x}^{\infty}\left(h(y)-\mathbb{E}h(Z)\right)e^{-y^2/2}\,dy\,, \end{aligned}$$

where we note that $\int_{-\infty}^{\infty}\left(h(y)-\mathbb{E}h(Z)\right)e^{-y^2/2}\,dy=0$.

The general form of the solution to our Stein equation is given by the particular solution f above plus some multiple (denoted by $ce^{-x^2/2}$) of the solution to the homogeneous equation.

On showing that our solution f is bounded, it therefore follows that this is the unique bounded solution of our Stein equation.

5. This follows from Theorem 3.4, our normal approximation result for sums of independent random variables, when we make the following two observations:

   (i) The function $h(x)=|x|$ is Lipschitz with Lipschitz constant 1, and therefore $h \in \mathcal{H}_W$.

   (ii) $\mathbb{E}|Z|=\sqrt{\frac{2}{\pi}}$, where $Z \sim \mathrm{N}(0,1)$. To see this, note that

$$\mathbb{E}|Z|=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}|x|e^{-x^2/2}\,dx=\sqrt{\frac{2}{\pi}}\int_{0}^{\infty}xe^{-x^2/2}\,dx=\sqrt{\frac{2}{\pi}}\,.$$

   From these two observations we have that $\left|\mathbb{E}|W|-\sqrt{\frac{2}{\pi}}\right| \leq d_W(W,Z)$, and the upper bound follows.