Online Appendix to Uncertainty in the Hot Hand Fallacy: Detecting Streaky Alternatives in Random Bernoulli Sequences

David M. Ritzwoller, Stanford University*
Joseph P. Romano, Stanford University

August 22, 2019

Contents

*E-mail: ritzwoll@stanford.edu, romano@stanford.edu.

A Supplementary Tables

List of Tables

Shooter k	1	2	3	4
101	0.8344	0.4824	0.2826	0.4393
102	0.4564	0.6254	0.8966	–
103	0.0234	0.5820	0.4895	0.7063
104	0.6944	0.8675	0.5782	0.9637
105	0.8367	0.8649	0.9052	0.9921
106	0.0347	0.0259	0.0061	0.0304
107	0.0531	0.1548	0.0214	0.0752
108	0.5563	0.1700	0.2443	0.0341
109	0.0001	0.0000	0.0008	0.0188
110	0.6755	0.1905	0.3701	0.4101
111	0.3244	0.4094	0.3053	0.3862
112	0.6369	0.5446	0.3575	0.2956
113	0.7322	0.5718	0.2251	0.5630
114	0.5514	0.5216	0.2222	0.4174
201	0.8697	0.8120	0.8690	0.8977
202	0.1131	0.4383	0.4165	0.2789
203	0.0441	0.3112	0.1748	0.0646
204	0.6336	0.4240	0.2888	0.4126
205	0.3457	0.1255	0.4082	0.4652
206	0.9026	0.8165	0.7996	0.8777
207	0.0758	0.0014	0.0367	0.1447
208	0.7059	0.3595	0.2742	0.0546
209	0.6996	0.6436	0.3159	0.5357
210	0.4051	0.0043	0.0055	0.0068
211	0.2857	0.6081	0.5949	0.4421
212	0.6834	0.9557	–	–

Online Appendix Table 1: Individual Permutation Test p-values: $\hat{D}_{n,k}(\mathbf{X}_i)$

Notes: Table displays the p-values for the individual level permutation tests rejecting for large values of $\hat{D}_{n,k}(\mathbf{X}_i)$. Each individual's shooting sequence is permuted $100,000$ times. $\hat{D}_{n,k}(\mathbf{X}_i)$ is computed on each permutation. The p-values are the proportions of permutations with $\hat{D}_{n,k}(\mathbf{X}_i)$ greater than or equal to the observed $\hat{D}_{n,k}(\mathbf{X}_i)$ among permutations where the statistic is defined.

Shooter k	1	2	3	4
101	0.8023	0.7377	0.5071	0.4210
102	0.3954	0.6891	0.6589	–
103	0.0301	0.2987	0.4161	0.3924
104	0.6551	0.8638	0.4610	0.5191
105	0.8083	0.4928	0.5365	0.6016
106	0.0203	0.1901	0.1451	0.3102
107	0.0309	0.1550	0.1606	0.4302
108	0.6110	0.0885	0.2356	0.0461
109	0.0000	0.0001	0.0003	0.0115
110	0.5804	0.4318	0.5372	0.4515
111	0.3629	0.2377	0.2791	0.3137
112	0.5676	0.6121	0.4195	0.4562
113	0.6758	0.7232	0.7916	0.8964
114	0.5514	0.4434	0.3773	0.4191
201	0.8452	0.5693	0.7285	0.7942
202	0.0953	0.3849	0.2608	0.1019
203	0.0360	0.3315	0.1475	0.1441
204	0.5771	0.3480	0.2910	0.3480
205	0.3457	0.0697	0.6178	0.4310
206	0.9026	0.6241	0.7670	0.7319
207	0.0635	0.0048	0.0729	0.1656
208	0.6645	0.2271	0.0160	0.1744
209	0.6996	0.3754	0.2401	0.6096
210	0.3733	0.0055	0.0145	0.0106
211	0.2196	0.5480	0.8024	0.3941
212	0.6447	0.7288	–	–

Online Appendix Table 2: Individual Permutation Test p-values: $\hat{P}_{n,k}(\mathbf{X}_i)$

Notes: Table displays the p-values for the individual level permutation tests rejecting for large values of $\hat{P}_{n,k}(\mathbf{X}_i)$. Each individual's shooting sequence is permuted $100,000$ times. $\hat{P}_{n,k}(\mathbf{X}_i)$ is computed on each permutation. The p-values are the proportions of permutations with $\hat{P}_{n,k}(\mathbf{X}_i)$ greater than or equal to the observed $\hat{P}_{n,k}(\mathbf{X}_i)$ among permutations where the statistic is defined.

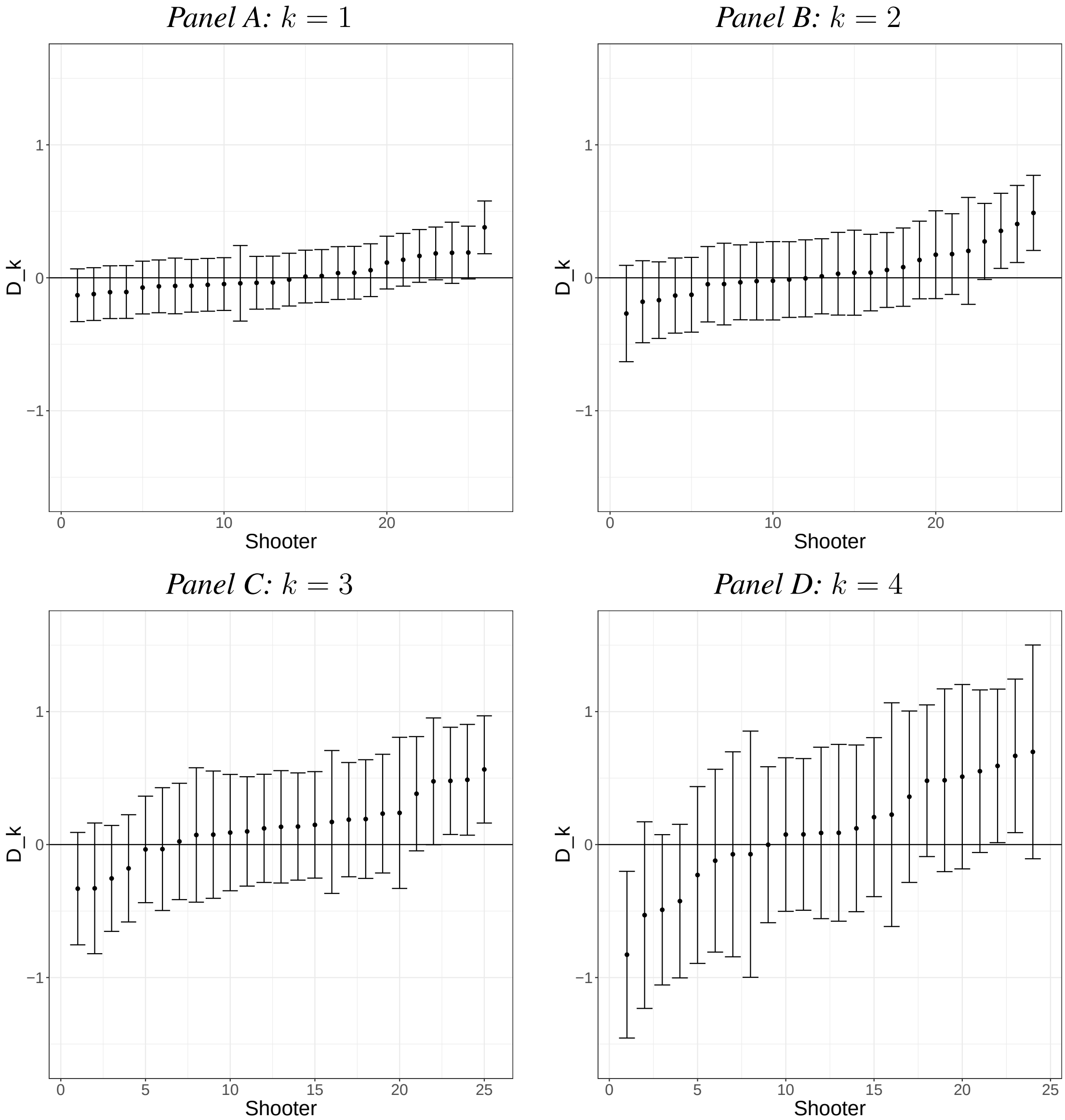
Shooter	k	1	2	3	4
101		0.8570	0.24453	0.1820	0.4406
102		0.4564	0.31095	0.7727	0.5004
103		0.0200	0.69354	0.4750	0.5406
104		0.7158	0.68479	0.6520	0.9251
105		0.8520	0.98531	0.9911	0.9831
106		0.0347	0.01722	0.0053	0.0185
107		0.0531	0.20156	0.0247	0.0271
108		0.4980	0.3553	0.3082	0.1010
109		0.0001	0.00498	0.0576	0.1506
110		0.6755	0.11305	0.3159	0.3797
111		0.2991	0.53313	0.3644	0.4169
112		0.6369	0.39707	0.3213	0.0994
113		0.7625	0.44023	0.0690	0.3680
114		0.4828	0.54219	0.1749	0.3998
201		0.8867	0.90162	0.8457	0.6846
202		0.1253	0.59909	0.8141	0.7434
203		0.0518	0.34993	0.2959	0.0122
204		0.6336	0.52413	0.1676	0.0902
205		0.2497	0.40972	0.0184	0.0730
206		0.8592	0.89041	0.6944	0.6613
207		0.0878	0.01222	0.1316	0.1631
208		0.7390	0.52155	0.6980	0.0717
209		0.6168	0.8569	0.4494	0.3422
210		0.4448	0.04747	0.0620	0.1018
211		0.2857	0.57369	0.2847	0.4637
212		0.6929	0.92258	0.8691	0.5952

Online Appendix Table 3: Individual Permutation Test p-values: $\hat{Q}_{n,k}(\mathbf{X}_i)$

Notes: Table displays the p-values for the individual level permutation tests rejecting for large values of $\hat{Q}_{n,k}(\mathbf{X}_i)$. Each individual's shooting sequence is permuted $100,000$ times. $\hat{Q}_{n,k}(\mathbf{X}_i)$ is computed on each permutation. The p-values are the proportions of permutations with $\hat{Q}_{n,k}(\mathbf{X}_i)$ greater than or equal to the observed $\hat{Q}_{n,k}(\mathbf{X}_i)$ among permutations where the statistic is defined.

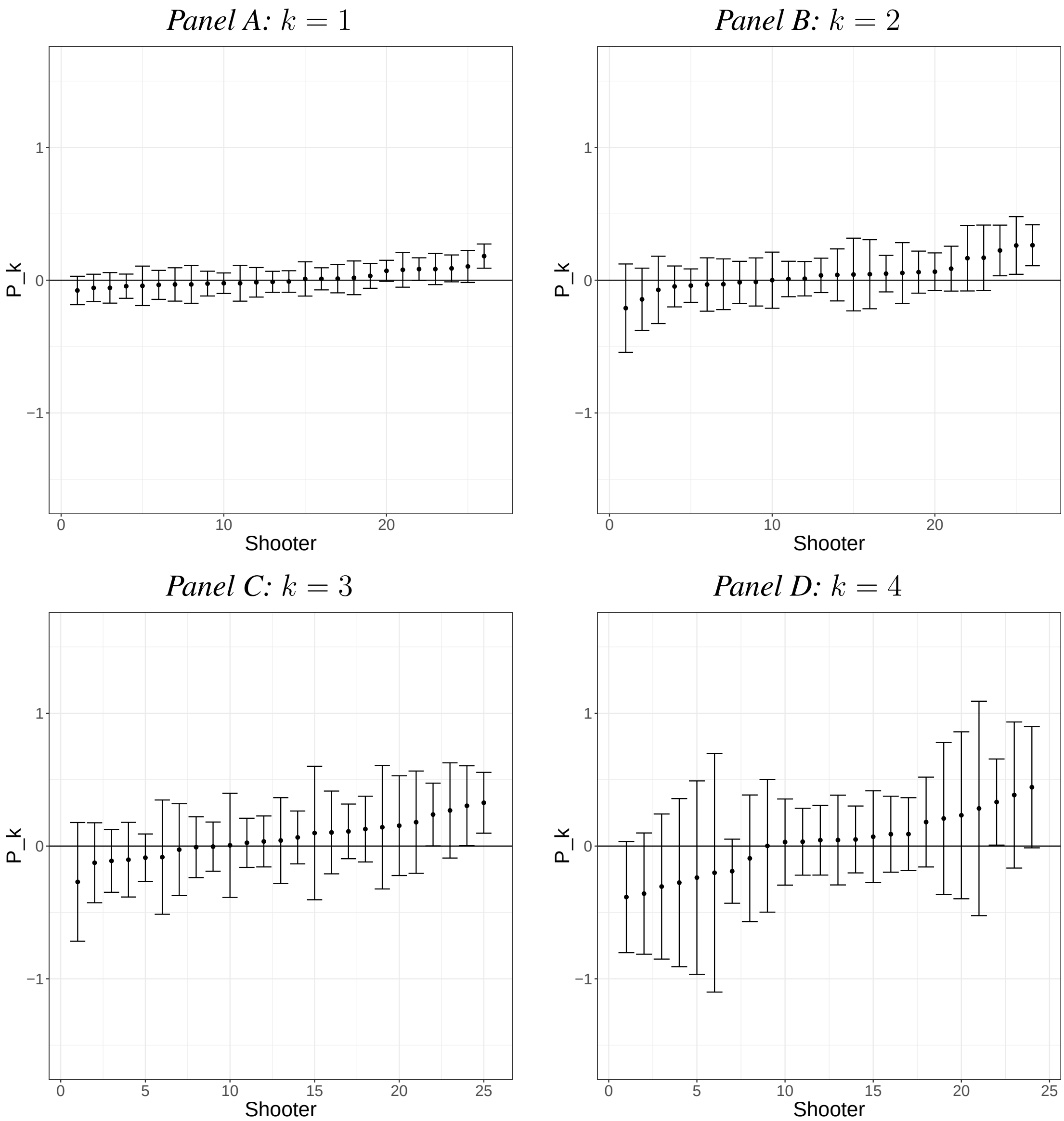
B Supplementary Figures

List of Figures



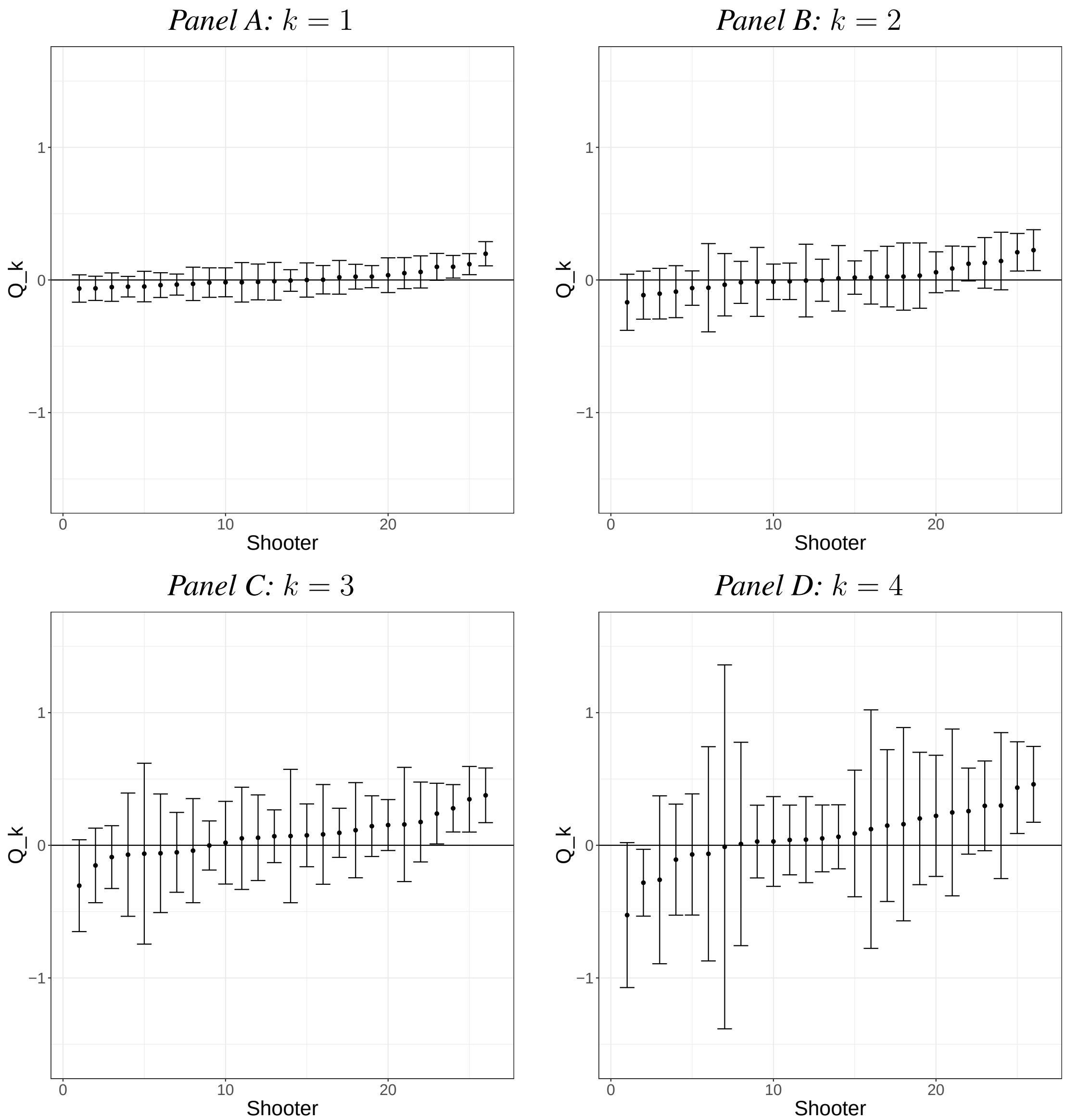
Online Appendix Figure 1: Normal Confidence Intervals: $\gamma_D(\mathbb{P}, k)$

Notes: Figure displays 95% confidence intervals for $\gamma_D(\mathbb{P}, k)$ for each k in $1, \ldots, 4$ and each shooter i with $\hat{D}_{n,k}(\mathbf{X}_i)$ defined. The 95% confidence intervals are given by $\hat{D}_{n,k}(\mathbf{X}_i) - \beta_D(n, k, \hat{p}_i) \pm t_{n,1-\alpha/2}\left(n_i^{-1/2}\sigma_D(\hat{p}_i, k)\right)$ where $\beta_D(n, k, \hat{p}_i)$ is the mean $\hat{D}_{n,k}(\mathbf{X}_i)$ computed for 100,000 bootstrap replicates of $\mathbf{X}_i$, $\hat{p}_i$ is the observed shooting percentage for shooter i, and n is the number of shots we observe for shooter i. Within each panel, we sort the shooters by $\hat{D}_{n,k}(\mathbf{X}_i) - \beta_D(n, k, \hat{p}_i)$, with the smallest value on the left and the largest value on the right.



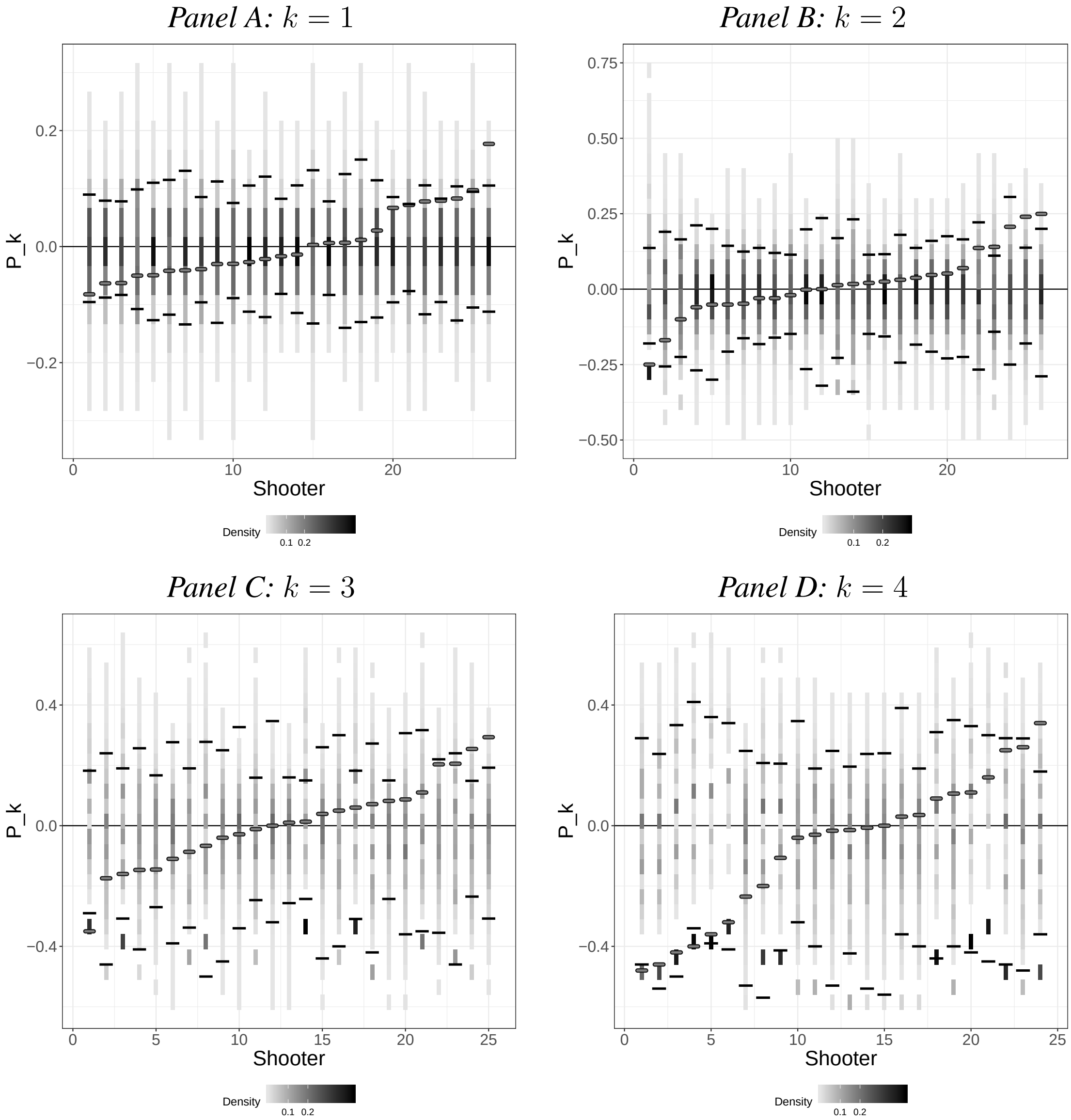
Online Appendix Figure 2: Normal Confidence Intervals: $\gamma_{\tilde{P}}\left(\mathbb{P},k\right)$

Notes: Figure displays 95% confidence intervals for $\gamma_{\tilde{P}}\left(\mathbb{P},k\right)$ for each k in $1,\ldots,4$ and each shooter i with $\hat{P}_{n,k}(\mathbf{X}_i)$ defined. The 95% confidence intervals are given by $\hat{P}_{n,k}\left(\mathbf{X}_i\right)-\hat{p}_i-\beta_P\left(n,k,\hat{p}_i\right) \pm t_{n,1-\alpha/2}\left(n_i^{-1/2}\sigma_{\tilde{P}}\left(\hat{p}_i,k\right)\right)$, where $\beta_P\left(n,k,\hat{p}_i\right)$ is the mean $\hat{P}_{n,k}(\mathbf{X}_i)-\hat{p}_i$ computed for $100,000$ bootstrap replicates of $\mathbf{X}_i$, $\hat{p}_i$ is the observed shooting percentage for shooter i, and n is the number of shots we observe for shooter i. Within each panel, we sort the shooters by $\hat{P}_{n,k}\left(\mathbf{X}_i\right)-\beta_P\left(n,k,\hat{p}_i\right)-\hat{p}_i$, with the smallest value on the left and the largest value on the right.



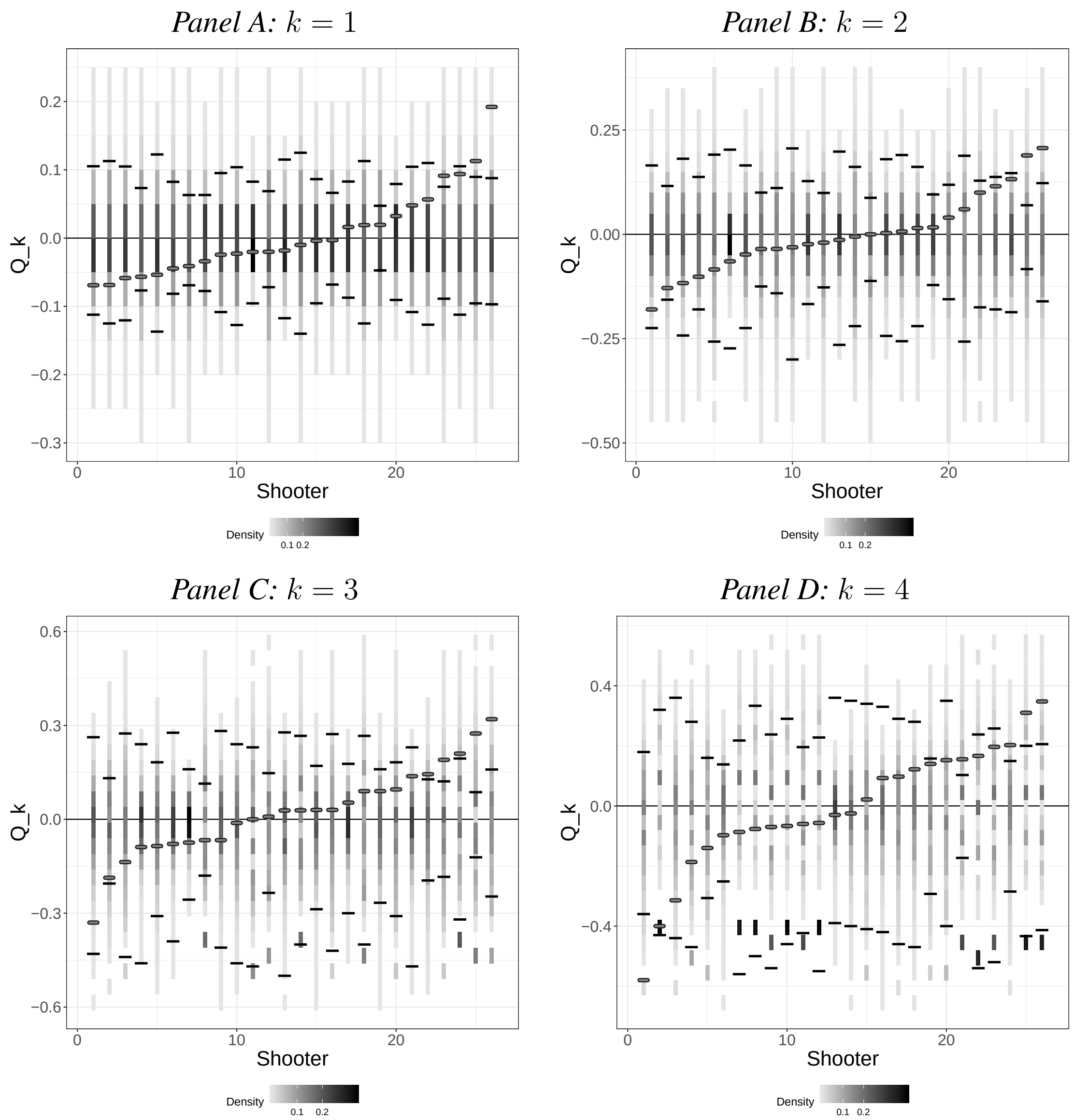
Online Appendix Figure 3: Normal Confidence Intervals: $\gamma_{\tilde{Q}}(\mathbb{P}, k)$

Notes: Figure displays 95% confidence intervals for $\gamma_{\tilde{Q}}(\mathbb{P}, k)$ for each k in $1, \ldots, 4$ and each shooter i with $\hat{Q}_{n,k}(\mathbf{X}_i)$ defined. The 95% confidence intervals are given by $\hat{Q}_{n,k}(\mathbf{X}_i) - (1 - \hat{p}_i) - \beta_Q(n, k, \hat{p}_i) \pm t_{n,1-\alpha/2}\left(n_i^{-1/2}\sigma_{\tilde{Q}}(1 - \hat{p}_i, k)\right)$ where $\beta_Q(n, k, \hat{p}_i)$ is the mean $\hat{Q}_{n,k}(\mathbf{X}_i) - (1 - \hat{p}_i)$ computed for $100,000$ bootstrap replicates of $\mathbf{X}_i$, $\hat{p}_i$ is the observed shooting percentage for shooter i, and n is the number of shots we observe for shooter i. Within each panel, we sort the shooters by $\hat{Q}_{n,k}(\mathbf{X}_i) - \beta_Q(n, k, \hat{p}_i) - (1 - \hat{p}_i)$, with the smallest value on the left and the largest value on the right.



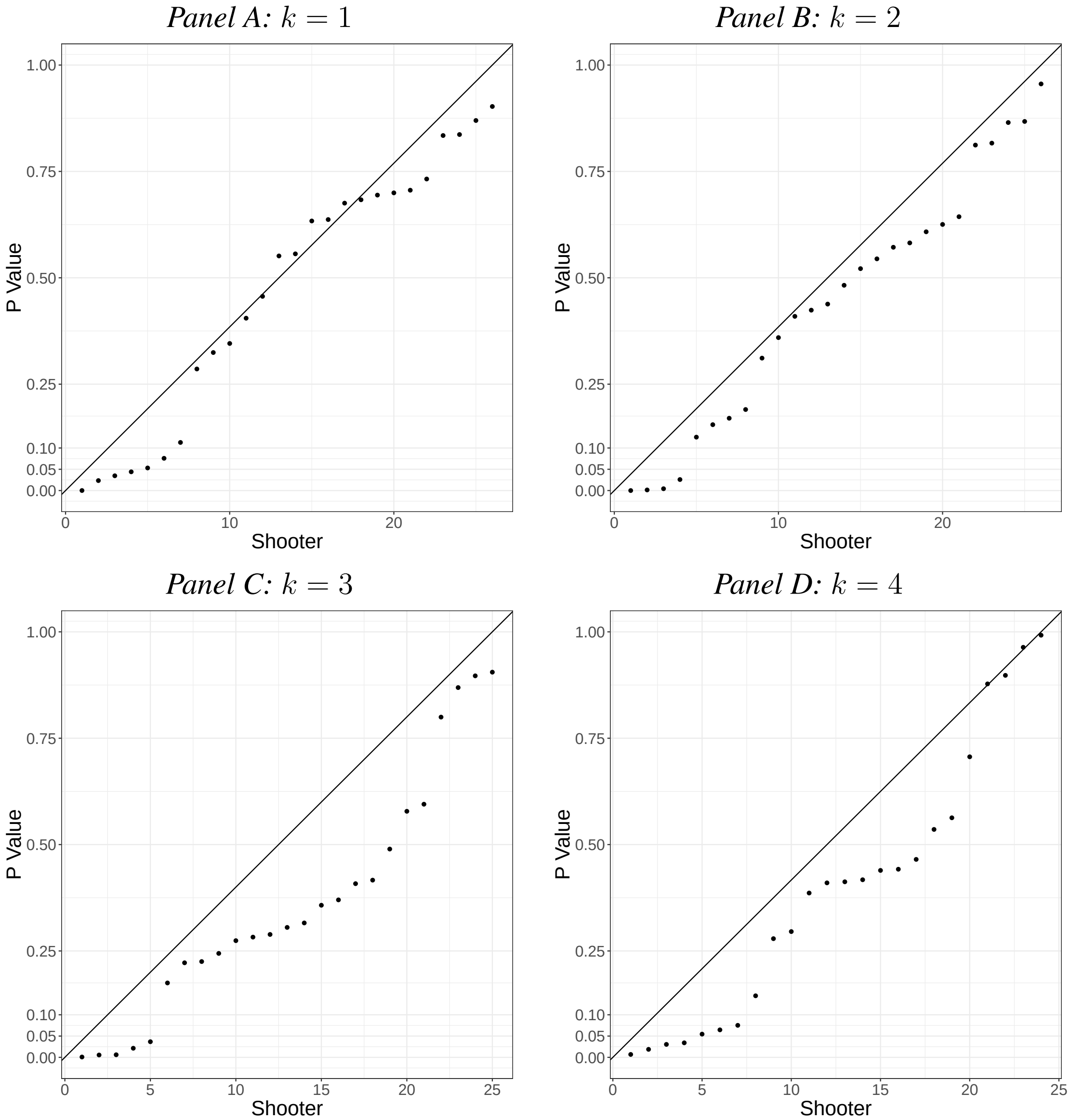
Online Appendix Figure 4: Permutation Distributions and Critical Values: $\hat{P}_{n,k}(\mathbf{X}_i) - \hat{p}_i$

Notes: Figure displays the observed values of $\hat{P}_{n,k}(\mathbf{X}_i) - \hat{p}_i$ overlaid onto the estimated permutation distribution of $\hat{P}_{n,k}(\mathbf{X}_i) - \hat{p}_i$ under H_0^i for each k in $1, \ldots, 4$ and each shooter i with $\hat{P}_{n,k}(\mathbf{X}_i)$ defined. The observed values of $\hat{P}_{n,k}(\mathbf{X}_i) - \hat{p}_i$ are denoted by light grey horizontal line segments. The estimated of the 97.5^{th} and 2.5^{th} quantiles of the permutation distribution of $\hat{P}_{n,k}(\mathbf{X}_i) - \hat{p}_i$ under H_0^i are denoted by black horizontal line segments. We estimate the permutation distribution of $\hat{P}_{n,k}(\mathbf{X}_i) - \hat{p}_i$ under H_0^i by permuting $\mathbf{X}_i$ $100,000$ times, computing $\hat{P}_{n,k}(\mathbf{X}_i) - \hat{p}_i$ for each permutation distribution. The estimates of the permutation distribution are displayed in vertical white to black gradients, shaded by the proportion of permutations whose computed value of $\hat{P}_{n,k}(\mathbf{X}_i) - \hat{p}_i$ lie in a fine partition of the observed support of $\hat{P}_{n,k}(\mathbf{X}_i) - \hat{p}_i$ under H_0. Within each panel, we sort the shooters by $\hat{P}_{n,k}(\mathbf{X}_i) - \hat{p}_i$, with the smallest value on the left and the largest value on the right.



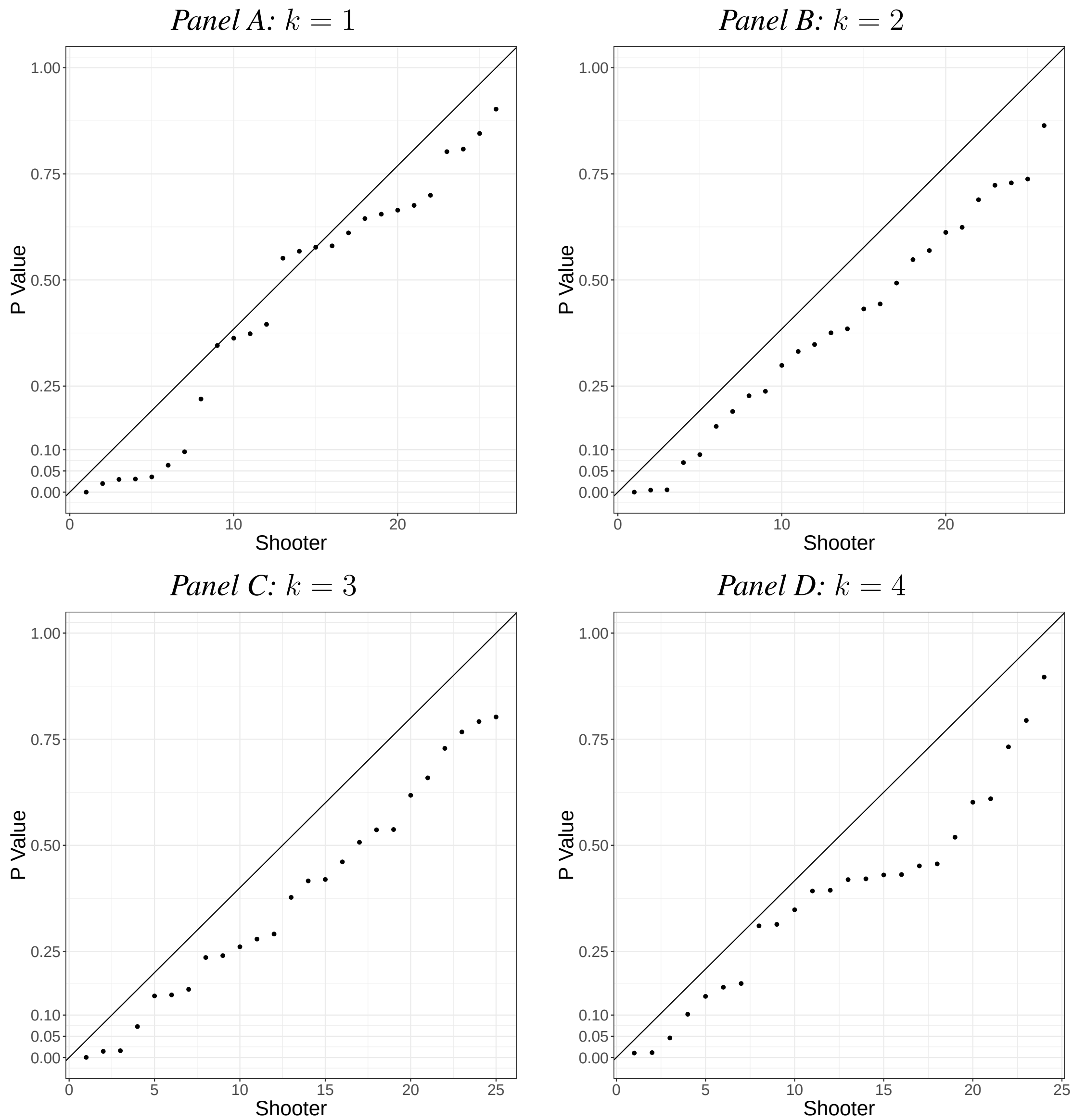
Online Appendix Figure 5: Permutation Distributions and Critical Values: $\hat{Q}_{n,k}(\mathbf{X}_i) - (1 - \hat{p}_i)$

Notes: Figure displays the observed values of $\hat{Q}_{n,k}(\mathbf{X}_i) - (1 - \hat{p}_i)$ overlaid onto the estimated permutation distribution of $\hat{Q}_{n,k}(\mathbf{X}_i) - (1 - \hat{p}_i)$ under H_0^i for each k in $1, \ldots, 4$ and each shooter i with $\hat{Q}_{n,k}(\mathbf{X}_i)$ defined. The observed values of $\hat{Q}_{n,k}(\mathbf{X}_i) - (1 - \hat{p}_i)$ are denoted by light grey horizontal line segments. The estimated of the 97.5^{th} and 2.5^{th} quantiles of the permutation distribution of $\hat{Q}_{n,k}(\mathbf{X}_i) - (1 - \hat{p}_i)$ under H_0^i are denoted by black horizontal line segments. We estimate the permutation distribution of $\hat{Q}_{n,k}(\mathbf{X}_i) - (1 - \hat{p}_i)$ under H_0^i by permuting $\mathbf{X}_i$ $100,000$ times, computing $\hat{Q}_{n,k}(\mathbf{X}_i) - (1 - \hat{p}_i)$ for each permutation distribution. The estimates of the permutation distribution are displayed in vertical white to black gradients, shaded by the proportion of permutations whose computed value of $\hat{Q}_{n,k}(\mathbf{X}_i) - (1 - \hat{p}_i)$ lie in a fine partition of the observed support of $\hat{Q}_{n,k}(\mathbf{X}_i) - (1 - \hat{p}_i)$ under H_0. Within each panel, we sort the shooters by $\hat{Q}_{n,k}(\mathbf{X}_i) - (1 - \hat{p}_i)$, with the smallest value on the left and the largest value on the right.



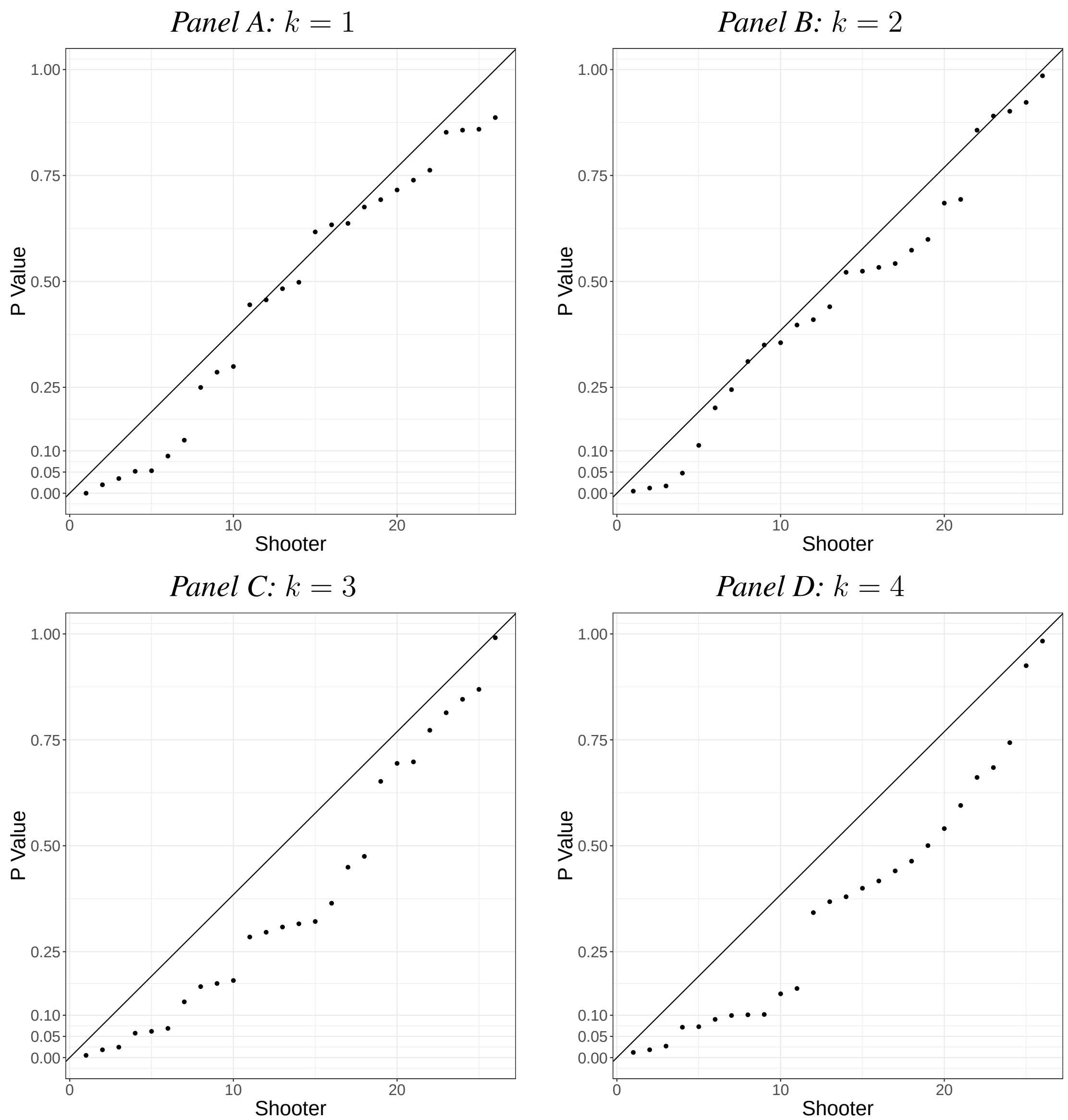
Online Appendix Figure 6: Individual Shooter Permutation Test p-values: $\hat{D}_{n,k}(\mathbf{X}_i)$

Notes: Figure displays the proportion of permutations of $\mathbf{X}_i$ with $\hat{D}_{n,k}(\cdot)$ greater than or equal to the observed $\hat{D}_{n,k}(\mathbf{X}_i)$ for each k in $1, \ldots, 4$ and each shooter i with $\hat{D}_{n,k}(\mathbf{X}_i)$ defined. Within each panel, we sort the p-values, with the smallest on the left and the largest on the right. The p-values are denoted by black dots.



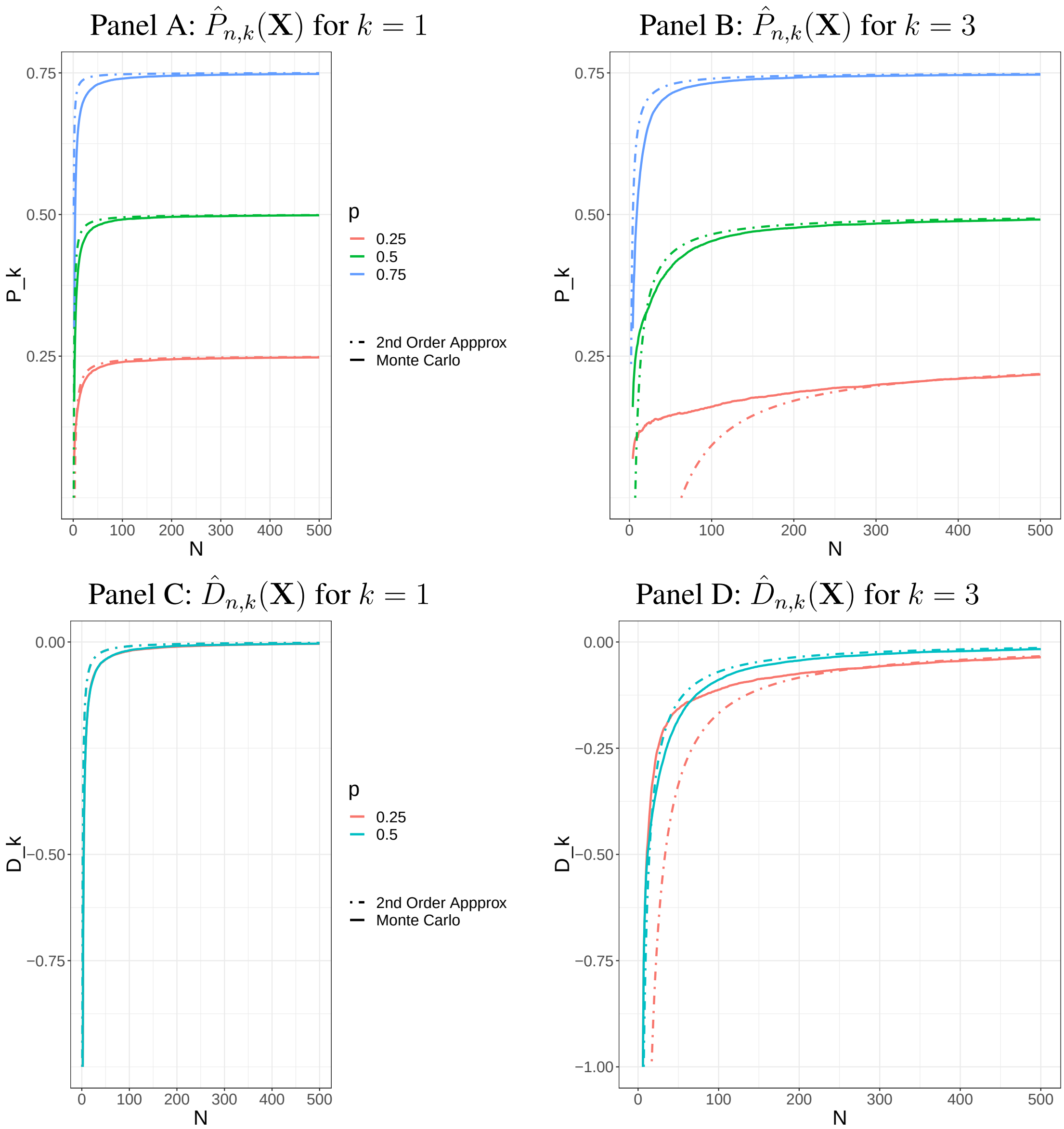
Online Appendix Figure 7: Individual Shooter Permutation Test p-values: $\hat{P}_{n,k}(\mathbf{X}_i)$

Notes: Figure displays the proportion of permutations of $\mathbf{X}_i$ with $\hat{P}_{n,k}(\cdot)$ greater than or equal to the observed $\hat{P}_{n,k}(\mathbf{X}_i)$ for each k in $1, \ldots, 4$ and each shooter i with $\hat{P}_{n,k}(\mathbf{X}_i)$ defined. Within each panel, we sort the p-values, with the smallest on the left and the largest on the right. The p-values are denoted by black dots.



Online Appendix Figure 8: Individual Shooter Permutation Test p-values: $\hat{Q}_{n,k}(\mathbf{X}_i)$

Notes: Figure displays the proportion of permutations of $\mathbf{X}_i$ with $\hat{Q}_{n,k}(\cdot)$ greater than or equal to the observed $\hat{Q}_{n,k}(\mathbf{X}_i)$ for each k in $1, \ldots, 4$ and each shooter i with $\hat{Q}_{n,k}(\mathbf{X}_i)$ defined. Within each panel, we sort the p-values, with the smallest on the left and the largest on the right. The p-values are denoted by black dots.



Online Appendix Figure 9: Second Order Approximation

Notes: The figure displays the the second order approximations to the expectations of $\hat{P}_{n,k}(\mathbf{X}_i)$ and $\hat{D}_{n,k}(\mathbf{X}_i)$ under H_0^i for $k = 1$ and 3 and $p \in (0.25, 0.5, 0.75)$. In all panels, the solid lines gives the Monte Carlo approximations and the dot-dashed lines give the second order approximation. Panels A and B display the Monte Carlo and second order approximations to $\hat{P}_{n,k}(\mathbf{X})$ for $k = 1$ and $k = 3$. Panels C and D display the Monte Carlo and second order approximations to $\hat{D}_{n,k}(\mathbf{X})$ for $k = 1$ and $k = 3$.

C Proofs for Theorems Stated in the Main Text

C.1 Proof of Theorem 2.1 (i)

Recall that $Y_{i,k} = \prod_{j=i}^{i+k} X_j$ and that $\hat{P}_{n,k}(\mathbf{X}) = \frac{\sum_{j=1}^{n-k} Y_{j,k}}{\sum_{j=1}^{n-k} Y_{j,k-1}}$. Note that $Y_{i,k}$ is k-dependent and strictly stationary. In order find the joint limiting distribution of $\left(Y_{ik}, Y_{i(k-1)}\right)$, we need to compute the asymptotic expectations, variances, and covariances for each of the terms. First, the expectations. We can see that $\mathbb{E}\left[Y_{j,k}\right] = p^{k+1}$. Next, the variances. Note that

$$
\begin{aligned}
\operatorname{Cov}\left(Y_{i,k}, Y_{i+u,k}\right) &= \mathbb{E}\left[Y_{i,k} Y_{i+u,k}\right] - \mathbb{E}\left[Y_{i,k}\right]\mathbb{E}\left[Y_{i+u,k}\right] \\
&= p^{k+1+|u|} - p^{2k+2}
\end{aligned}
$$

for $|u| \leq k$. Therefore,

$$
\begin{aligned}
\sum_{u=-k}^{u=k} \operatorname{Cov}\left(Y_{i,k}, Y_{i+u,k}\right) &= \sum_{u=-k}^{u=k}\left(p^{k+1+|u|} - p^{2k+2}\right) \\
&= p^{k+1}\left(1 - p^{k+1}\right) + 2\sum_{u=1}^{k}\left(p^{k+1+u} - p^{2k+2}\right) \\
&= p^{k+1} - p^{2k+2} + 2\sum_{u=1}^{k}\left(p^{k+1+u}\right) - 2kp^{2k+2} \\
&= p^{k+1} - (2k+1)\,p^{2k+2} + 2p^{k+1}\left(\frac{p\left(1-p^{k}\right)}{1-p}\right) \\
&= p^{k+1} - (2k+1)\,p^{2k+2} + \frac{2p^{k+2} - 2p^{2k+2}}{1-p}.
\end{aligned} \tag{C.1}
$$

Next, we compute the covariance. Note that

$$
\begin{aligned}
\operatorname{Cov}\left(Y_{i,k}, Y_{i+u,k-1}\right) &= \mathbb{E}\left[Y_{i,k} Y_{i+u,k-1}\right] - \mathbb{E}\left[Y_{i,k}\right]\mathbb{E}\left[Y_{i+u,k-1}\right] \\
&= \mathbb{E}\left[\prod_{j=i}^{i+k} X_j \prod_{j=i+u}^{i+u+k-1} X_j\right] - p^{2k+1} \\
&= \begin{cases}
p^{k+u} - p^{2k+1} & \text{if } 1 < u \leq k \\
p^{k+1} - p^{2k+1} & \text{if } u \in \{0,1\} \\
p^{k+1+|u|} - p^{2k+1} & \text{if } -k < u < 0 \\
0 & \text{if } u \leq -k \text{ or } u > k.
\end{cases}
\end{aligned}
$$

Therefore,

$$\begin{aligned}\sum_{u=-k}^{u=k}\mathrm{Cov}\left(Y_{i,k},Y_{i+u,k-1}\right)&=2p^{k+1}-2p^{2k+1}+\sum_{u=2}^{k}\left(p^{k+u}-p^{2k+1}\right)+\sum_{u=1}^{k-1}\left(p^{k+1+u}-p^{2k+1}\right)\\&=2p^{k+1}-2kp^{2k+1}+\sum_{u=2}^{k}p^{k+u}+\sum_{u=1}^{k-1}p^{k+1+u}\\&=2p^{k+1}-2kp^{2k+1}+2p^{k+1}\sum_{u=1}^{k-1}p^{u}\\&=2p^{k+1}-2kp^{2k+1}+2p^{k+1}\left(\frac{p\left(1-p^{k-1}\right)}{1-p}\right)\\&=2p^{k+1}-2kp^{2k+1}+\frac{2p^{k+2}-2p^{2k+1}}{1-p}.\end{aligned}\tag{C.2}$$

By Theorem 1.7 of Ibragimov (1962) and the Cramér-Wold device, we have that

$$\begin{aligned}&(n-k)^{1/2}\left(\left(\sum_{i=1}^{n-k}Y_{i,k},\sum_{i=1}^{n-k}Y_{i,k-1}\right)^{\top}-\left(p^{k+1},p^{k}\right)^{\top}\right)\\&\xrightarrow{d}N\left(0,\begin{bmatrix}p^{k+1}-(2k+1)\,p^{2k+2}+\frac{2p^{k+2}-2p^{2k+2}}{1-p} & 2p^{k+1}-2kp^{2k+1}+\frac{2p^{k+2}-2p^{2k+1}}{1-p}\\ 2p^{k+1}-2kp^{2k+1}+\frac{2p^{k+2}-2p^{2k+1}}{1-p} & p^{k}-(2k-1)\,p^{2k}+\frac{2p^{k+1}-2p^{2k}}{1-p}\end{bmatrix}\right).\end{aligned}$$

Next, we use the delta method to evaluate the asymptotic distribution of $\hat{P}_{n,k}\left(\mathbf{X}\right)$. Let $g(\theta_1,\theta_2)=\theta_1/\theta_2$. Then $\nabla g\left(\theta_1,\theta_2\right)=\left(\frac{1}{\theta_2},\frac{-\theta_1}{\theta_2^2}\right)^{\top}$. Evaluating at $\left(p^{k+1},p^{k}\right)^{\top}$, we find

$$\nabla g\left(p^{k+1},p^{k}\right)=\begin{bmatrix}p^{-k}\\-p^{1-k}\end{bmatrix}.$$

Note that

$$\begin{aligned}&\begin{bmatrix}p^{-k} & -p^{1-k}\end{bmatrix}\begin{bmatrix}p^{k+1}-(2k+1)\,p^{2k+2}+\frac{2p^{k+2}-2p^{2k+2}}{1-p} & 2p^{k+1}-2kp^{2k+1}+\frac{2p^{k+2}-2p^{2k+1}}{1-p}\\ 2p^{k+1}-2kp^{2k+1}+\frac{2p^{k+2}-2p^{2k+1}}{1-p} & p^{k}-(2k-1)\,p^{2k}+\frac{2p^{k+1}-2p^{2k}}{1-p}\end{bmatrix}\begin{bmatrix}p^{-k}\\-p^{1-k}\end{bmatrix}\\&=\begin{bmatrix}p(1-p^{k+1})\\p(1-p^{k})\end{bmatrix}^{\top}\begin{bmatrix}p^{-k}\\-p^{1-k}\end{bmatrix}\\&=p^{1-k}(1-p).\end{aligned}$$

Hence, we have that

$$n^{1/2}\left(\hat{P}_{n,k}\left(\mathbf{X}\right)-p\right)\xrightarrow{d}N(0,p^{1-k}\left(1-p\right)).$$

C.2 Proof of Theorem 2.1 (ii)

We first find the joint limiting distribution of $(Y_{i,k},Y_{i,k-1},X_i)$. We computed most of the necessary asymptotic expectations, variances, and covariances in the proof of Theorem 2.1 (i). Note that $\mathbb{E}\left[X_i\right]=p$, that $Var\left(X_i\right)=p\left(1-p\right)$, and that

$$\text{Cov}\left(Y_{i,k},X_{i+u}\right)=\begin{cases}p^{k+1}-p^{k+2} & \text{if } 0\leq u\leq k\\ 0 & \text{otherwise.}\end{cases}$$

Therefore, we can evaluate

$$\sum_{u=-k}^{u=k}\text{Cov}\left(Y_{i,k},X_{i+u}\right)=\left(k+1\right)\left(p^{k+1}-p^{k+2}\right)$$

and

$$\sum_{u=-k}^{u=k}\text{Cov}\left(Y_{i,k-1},X_{i+u}\right)=k\left(p^{k}-p^{k+1}\right).$$

By Theorem 1.7 of Ibragimov (1962) and the Cramér-Wold device, we have that

$$\left(n-k\right)^{1/2}\left(\left(\sum_{i=1}^{n-k}Y_{i,k},\sum_{i=1}^{n-k}Y_{i,k-1},\sum_{i=1}^{n-k}X_i\right)^{\top}-\left(p^{k+1},p^{k},p\right)^{\top}\right)\xrightarrow{d}N\left(0,V\right),$$

where

$$V=\left[\begin{array}{ccc}p^{k+1}-\left(2k+1\right)p^{2k+2}+\frac{2p^{k+2}-2p^{2k+2}}{1-p} & 2p^{k+1}-2kp^{2k+1}+\frac{2p^{k+2}-2p^{2k+1}}{1-p} & \left(k+1\right)\left(p^{k+1}-p^{k+2}\right)\\ 2p^{k+1}-2kp^{2k+1}+\frac{2p^{k+2}-2p^{2k+1}}{1-p} & p^{k}-\left(2k-1\right)p^{2k}+\frac{2p^{k+1}-2p^{2k}}{1-p} & k\left(p^{k}-p^{k+1}\right)\\ \left(k+1\right)\left(p^{k+1}-p^{k+2}\right) & k\left(p^{k}-p^{k+1}\right) & p\left(1-p\right)\end{array}\right].$$

Note that

$$n^{1/2}\left(n^{-1}\sum_{i=1}^{n}X_i\Big/\left(n-k\right)^{-1}\sum_{i=1}^{n}X_i-1\right)\xrightarrow{p}0,$$

so we can replace $\frac{1}{n-k}\sum_{i=1}^{n-k}X_i$ with $\frac{1}{n}\sum_{i=1}^{n}X_i$.

Next, use the delta method to evaluate the asymptotic distribution of $\hat{P}_{n,k}\left(\mathbf{X}\right)-\hat{p}$. Let $g(\theta_1,\theta_2,\theta_3)=$

$\theta_1/\theta_2 - \theta_3$. Then $\nabla g(\theta_1, \theta_2, \theta_3) = \left(\frac{1}{\theta_2}, \frac{-\theta_1}{\theta_2^2}, -1\right)^\top$. Evaluating at $\left(p^{k+1}, p^k, p\right)^\top$, we find

$$\nabla g\left(p^{k+1}, p^k, p\right) = \begin{bmatrix} p^{-k} \\ -p^{1-k} \\ -1 \end{bmatrix}.$$

Note that

$$\begin{aligned}
&\begin{bmatrix} p^{-k} \\ -p^{1-k} \\ -1 \end{bmatrix}^\top V \begin{bmatrix} p^{-k} \\ -p^{1-k} \\ -1 \end{bmatrix} \\
&= \begin{bmatrix} p - (1 + k(1-p))\, p^{1+k} \\ p - (k(1-p) + p)\, p^k \\ 0 \end{bmatrix}^\top \begin{bmatrix} p^{-k} \\ -p^{1-k} \\ -1 \end{bmatrix} \\
&= p^{1-k}(1-p)\left(1-p^k\right).
\end{aligned}$$

and that $p^{k+1}/p^k - p = 0$. Hence,

$$n^{1/2}\left(\hat{P}_{n,k}(\mathbf{X}) - \hat{p}\right) \xrightarrow{d} N\left(0, p^{1-k}(1-p)\left(1-p^k\right)\right).$$

C.3 Proof of Theorem 2.1 (iii)

Recall that $Z_{i,k} = \prod_{j=i}^{i+k}(1 - X_j)$, that

$$\hat{Q}_{n,k}(\mathbf{X}) = \frac{\sum_{i=1}^{n-k} Z_{k,i}}{\sum_{i=k}^{n-k} Z_{k-1,i}},$$

and that

$$\hat{D}_{n,k}(\mathbf{X}) = \frac{\sum_{i=1}^{n-k} Y_{k,i}}{\sum_{i=1}^{n-k} Y_{k-1,i}}. - \left(1 - \frac{\sum_{i=1}^{n-k} Z_{k,i}}{\sum_{i=k}^{n-k} Z_{k-1,i}}\right).$$

We first find the joint limiting distribution of $(Y_{k,i}, Y_{k-1,i}, Z_{k,i}, Z_{k-1,i})$. We need to compute the asymptotic expectations, variances, and covariances for each of the terms.

First, the expectations. We can see that $\mathbb{E}[Y_{j,k}] = p^{k+1}$ and $\mathbb{E}[Z_{j,k}] = (1-p)^{k+1}$. Next, the variances. Recall from the proof of Theorem 2.1 (i) that $\sum_{u=-k}^{u=k} \mathrm{Cov}(Y_{i,k}, Y_{i+u,k})$ is given by (C.1) and so therefore

$$\sum_{u=-k}^{u=k} \text{Cov}\left(Z_{i,k}, Z_{i+u,k}\right) = (1-p)^{k+1} - (2k+1)(1-p)^{2k+2} + \frac{2(1-p)^{k+2} - 2(1-p)^{2k+2}}{p}.$$

Next, compute the covariances. Recall from the proof of Theorem 2.1 (i) that $\sum_{u=-k}^{u=k} \text{Cov}\left(Y_{i,k}, Y_{i+u,k-1}\right)$ is given by (C.2) and so therefore

$$\sum_{u=-k}^{u=k} \text{Cov}\left(Z_{i,k}, Z_{i+u,k-1}\right) = 2(1-p)^{k+1} - 2k(1-p)^{2k+1} + \frac{2(1-p)^{k+2} - 2(1-p)^{2k+1}}{p}.$$

Note that

$$\begin{aligned}
\text{Cov}\left(Y_{i,k}, Z_{i+u,k}\right) &= \mathbb{E}\left[Y_{i,k} Z_{i+u,k}\right] - \mathbb{E}\left[Y_{i,k}\right]\mathbb{E}\left[Z_{i+u,k}\right] \\
&= \mathbb{E}\left[Y_{i,k} Z_{i+u,k}\right] - p^{k+1}(1-p)^{k+1} \\
&= -p^{k+1}(1-p)^{k+1} \text{ for } -k \leq u \leq k \\
\text{Cov}\left(Y_{i,k}, Z_{i+u,k-1}\right) &= \mathbb{E}\left[Y_{i,k} Z_{i+u,k-1}\right] - \mathbb{E}\left[Y_{i,k}\right]\mathbb{E}\left[Z_{i+u,k-1}\right] \\
&= \mathbb{E}\left[Y_{i,k} Z_{i+u,k-1}\right] - p^{k+1}(1-p)^{k} \\
&= -p^{k+1}(1-p)^{k} \text{ for } -k-1 \leq u \leq k \\
\text{Cov}\left(Y_{i,k-1}, Z_{i+u,k}\right) &= \mathbb{E}\left[Y_{i,k-1} Z_{i+u,k}\right] - \mathbb{E}\left[Y_{i,k-1}\right]\mathbb{E}\left[Z_{i+u,k}\right] \\
&= -p^{k}(1-p)^{k+1} \text{ for } -k \leq u \leq k-1
\end{aligned}$$

As $\mathbb{E}\left[Y_{i,k} Z_{i+u,k}\right]$, $\mathbb{E}\left[Y_{i,k} Z_{i+u,k-1}\right]$, and $\mathbb{E}\left[Y_{i,k-1} Z_{i+u,k}\right]$ are all equal to zero if the there is any overlap in the X_j's and $(1-X_k)$'s composing $Y_{i,r}$ and $Z_{i+u,s}$ for any r and s.We can see that

$$\begin{aligned}
\sum_{u=-k}^{y=k} \text{Cov}\left(Y_{i,k}, Z_{i+u,k}\right) &= -(2k+1)\, p^{k+1}(1-p)^{k+1} \\
\sum_{u=-k}^{y=k} \text{Cov}\left(Y_{i,k+1}, Z_{i+u,k-1}\right) &= -(2k)\, p^{k+1}(1-p)^{k} \\
\sum_{u=-k+1}^{y=k+1} \text{Cov}\left(Y_{i,k}, Z_{i+u,k+1}\right) &= -(2k)\, p^{k}(1-p)^{k+1}.
\end{aligned}$$

Therefore, by Theorem 1.7 of Ibragimov (1962) and the Cramér-Wold device, we have that

$$(n-k)^{1/2}\left(\left(\sum_{i=1}^{n-k} Y_{k,i}, \sum_{i=1}^{n-k} Y_{k-1,i}, \sum_{i=1}^{n-k} Z_{k,i}, \sum_{i=1}^{n-k} Z_{k-1,i}\right)^{\top} - \left(p^{k+1}, p^{k}, (1-p)^{k+1}, (1-p)^{k}\right)\right) \tag{C.3}$$

$$\xrightarrow{d} N\left(0, \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix}\right).$$

where

$$V_{11} = \begin{bmatrix} p^{k+1} - (2k+1)\,p^{2k+2} + \frac{2p^{k+2}-2p^{2k+2}}{1-p} & 2p^{k+1} - 2kp^{2k+1} + \frac{2p^{k+2}-2p^{2k+1}}{1-p} \\ 2p^{k+1} - 2kp^{2k+1} + \frac{2p^{k+2}-2p^{2k+1}}{1-p} & p^{k} - (2k-1)\,p^{2k} + \frac{2p^{k+1}-2p^{2k}}{1-p} \end{bmatrix},$$

$$V_{21} = \begin{bmatrix} -(2k+1)\,p^{k+1}(1-p)^{k+1} & -(2k)\,p^{k}(1-p)^{k+1} \\ -(2k)\,p^{k+1}(1-p)^{k} & -(2k-1)\,p^{k}(1-p)^{k} \end{bmatrix},$$

$$V_{12} = \begin{bmatrix} -(2k+1)\,p^{k+1}(1-p)^{k+1} & -(2k)\,p^{k+1}(1-p)^{k} \\ -(2k)\,p^{k}(1-p)^{k+1} & -(2k-1)\,p^{k}(1-p)^{k} \end{bmatrix}, \text{ and}$$

$$V_{22} = \begin{bmatrix} (1-p)^{k+1} - (2k+1)(1-p)^{2k+2} + \frac{2(1-p)^{k+2}-2(1-p)^{2k+2}}{p} & 2(1-p)^{k+1} - 2k(1-p)^{2k+1} + \frac{2(1-p)^{k+2}-2(1-p)^{2k+1}}{p} \\ 2(1-p)^{k+1} - 2k(1-p)^{2k+1} + \frac{2(1-p)^{k+2}-2(1-p)^{2k+1}}{p} & (1-p)^{k} - (2k-1)(1-p)^{2k} + \frac{2(1-p)^{k+1}-2(1-p)^{2k}}{p} \end{bmatrix}.$$

Next, use the delta method to evaluate the asymptotic distribution of $\hat{D}_{n,k}(\mathbf{X})$. Let $g((\theta_1,\theta_2,\theta_3,\theta_4)^{\top}) = \theta_1/\theta_2 + \theta_3/\theta_4 - 1$. Then $\nabla g\left((\theta_1,\theta_2,\theta_3,\theta_4)^{\top}\right) = \left(\frac{1}{\theta_2}, \frac{-\theta_1}{\theta_2^2}, \frac{1}{\theta_4}, \frac{-\theta_3}{\theta_4^2}\right)^{\top}$. Evaluating at $\left(p^{k+1}, p^{k}, (1-p)^{k+1}, (1-p)^{k}\right)^{\top}$, we find $g\left(\left(p^{k+1}, p^{k}, (1-p)^{k+1}, (1-p)\right)^{\top}\right) = 0$ and

$$\nabla g\left(\left(p^{k+1}, p^{k}, (1-p)^{k+1}, (1-p)\right)^{\top}\right) = \begin{bmatrix} p^{-k} \\ -p^{1-k} \\ (1-p)^{-k} \\ -(1-p)^{1-k} \end{bmatrix}.$$

Note that

$$\begin{bmatrix} p^{-k} \\ -p^{1-k} \\ (1-p)^{-k} \\ -(1-p)^{1-k} \end{bmatrix}^\top \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix} \begin{bmatrix} p^{-k} \\ -p^{1-k} \\ (1-p)^{-k} \\ -(1-p)^{1-k} \end{bmatrix}$$
$$= \begin{bmatrix} p\left(1-p^k\right) \\ p-p^k \\ (1-p)\left(1-(1-p)^k\right) \\ (1-p)-(1-p)^k \end{bmatrix}^\top \begin{bmatrix} p^{-k} \\ -p^{1-k} \\ (1-p)^{-k} \\ -(1-p)^{1-k} \end{bmatrix}$$
$$= (p(1-p))^{1-k}\left((1-p)^k + p^k\right)$$

Hence, we have that

$$n^{1/2}\hat{D}_{n,k}(\mathbf{X}) \xrightarrow{d} N\left(0, (p(1-p))^{1-k}\left((1-p)^k + p^k\right)\right).$$

C.4 Proof of Theorem 2.2 (i)

First we prove the result in the case $k = 1$. Recall that $V_{n,k} = \sum_{i=1}^{n-k} Y_{i,k}$ and $W_{n,k} = \sum_{i=1}^{n-k} Z_{i,k}$ where $Y_{i,k} = \prod_{j=i}^{i+k} X_i$ and $Z_{i,k} = \prod_{j=i}^{i+k}(1-X_i)$. The test statistic under consideration is $\sqrt{n}\hat{D}_{n,1}(\mathbf{X})$, where

$$\begin{aligned}\hat{D}_{n,1}(X_1,\ldots,X_n) &= \frac{V_{n,1}}{V_{n,0}} - \left(1 - \frac{W_{n,1}}{W_{n,0}}\right) \\ &= \frac{V_{n,1}}{V_{n,0}} - \frac{\sum_{i=1}^{n-1}(1-X_i)X_{i+1}}{n-V_{n,0}} = \frac{V_{n,1}}{V_{n,0}} - \frac{\sum_{i=2}^{n} X_i - V_{n,1}}{n-V_{n,0}} \\ &= \frac{\frac{V_{n,1}}{n} - \frac{\left(\sum_{i=2}^{n} X_i\right)\cdot V_{n,0}}{n^2}}{\frac{V_{n,0}}{n}\left(1-\frac{V_{n,0}}{n}\right)}.\end{aligned} \tag{C.4}$$

The denominator of (C.4) tends to $p(1-p)$ with probability one. Since $|V_{n,0} - \sum_{i=2}^{n} X_i| \leq 2$, it follows that

$$\sqrt{n}\left(\frac{V_{n,1}}{n} - \frac{\left(\sum_{i=2}^{n} X_i\right)\cdot V_{n,0}}{n^2}\right) - \sqrt{n}\left(\frac{V_{n,1}}{n} - \frac{V_{n,0}V_{n,0}}{n^2}\right) \xrightarrow{P} 0.$$

Therefore,

$$\sqrt{n}\hat{D}_{n,1}(\mathbf{X}) = \frac{\sqrt{n}\left(\frac{V_{n,1}}{n} - \frac{V_{n,0}^2}{n^2}\right)}{\frac{V_{n,0}}{n}\left(1 - \frac{V_{n,0}}{n}\right)} + o_P(1). \tag{C.5}$$

Let

$$\Pi = \Pi_n = (\Pi(1), \ldots, \Pi(n))$$

denote a (uniform) random permutation of $(1, 2, \ldots n)$ independent of X_i for all i.

We need to analyze the joint limiting distribution of $\sqrt{n}\left(\hat{D}_{n,1}(\mathbf{X}), \tilde{D}_{n,1}(\mathbf{X})\right)$, where

$$\tilde{D}_{n,1}(\mathbf{X}) = \hat{D}_{n,1}\left(X_{\Pi(1)}, \ldots, X_{\Pi(n)}\right).$$

In particular, to verify Hoeffding's condition (See Lehmann and Romano 2005, Theorem 15.2.3), we need to verify that $\sqrt{n}\hat{D}_{n,1}(\mathbf{X})$ and $\sqrt{n}\tilde{D}_{n,1}(\mathbf{X})$ are asymptotically independent. We already know that the marginal distributions of the joint limiting distribution of $\sqrt{n}\hat{D}_{n,1}(\mathbf{X})$ and $\sqrt{n}\tilde{D}_{n,1}(\mathbf{X})$ are standard normal. Since the denominator $V_{n,0}/n\,(1 - V_{n,0}/n)$ is invariant under permutations and tends to $p(1-p)$ with probability one, it suffices to show $\sqrt{n}f\left(\bar{V}_{n,1}, \bar{V}_{n,0}\right)$ and $\sqrt{n}f\left(\tilde{V}_{n,1}, \bar{V}_{n,0}\right)$ are asymptotically independent, where

$$f(v_1, v_0) = v_1 - v_0^2 \quad \bar{V}_{n,1} = V_{n,1}/n \quad \bar{V}_{n,0} = V_{n,0}/n$$

and

$$\tilde{V}_{n,1} = \frac{1}{n}\sum_{i=1}^{n-1} X_{\Pi(i)}X_{\Pi(i+1)}.$$

Note that $\sqrt{n}\left(\bar{V}_{n,1}, \bar{V}_{n,0}\right)$ and $\sqrt{n}\left(\tilde{V}_{n,1}, \bar{V}_{n,0}\right)$ are not asymptotically independent. Next apply the Delta Method, noting that averaging over n or $n-1$ does not affect our analysis. So,

$$\sqrt{n}f\left(\bar{V}_{n,1}, \bar{V}_{n,0}\right) = \sqrt{n}\left[\bar{V}_{n,1} - p^2 - 2p\left(\bar{V}_{n,0} - p\right)\right] + o_p(1).$$

The variance of the linear approximation is equal to

$$\begin{aligned} &n\left[\operatorname{Var}\left(\bar{V}_{n,1}\right) + 2p^2\operatorname{Var}\left(\bar{V}_{n,0}\right) - 4p\operatorname{Cov}\left(\bar{V}_{n,1}, \bar{V}_{n,0}\right)\right] \\ &= p^2\left(1-p^2\right) + 4\left(p^3 - p^4\right) + 4p^3(1-p) - 8p\left(p^2 - p^3\right) = p^2(1-p)^2. \end{aligned}$$

Using the linear approximation for both the original statistic and the permuted statistic, we can apply the Cramér-Wold device. Let

$$T_i = n^{-1/2}\left\{a\left[X_iX_{i+1} - p^2 - 2p\left(X_i - p\right)\right] + b\left[X_{\Pi(i)}X_{\Pi(i+1)} - p^2 - 2p\left(X_{\Pi(i)} - p\right)\right]\right\}. \quad \text{(C.6)}$$

Therefore, it suffices to show that, for any a and b, $\sum_{i=1}^n T_i$ is asymptotically normal with mean 0 and variance $\left(a^2+b^2\right)p^2\left(1-p\right)^2$.

We now determine the limiting behavior of $\sum_i T_i$ by applying a Central Limit Theorem of Stein (1986) for dependent random variables in a form stated in Rinott (1994). To do this, we will condition on $\Pi_n = \pi_n$, so that Theorem 2.2 of Rinott (1994) is applicable. For shorthand, drop the subscript on $\pi_n = \pi$. Let S_i be the set of indices j such that T_i and T_j are dependent. Clearly, S_i contains both i and $i+1$. But, S_i also contains the indices j for which $\pi\left(j\right) = i$ or $\pi\left(j\right) = i+1$, as well as the indices j for which $\pi\left(j+1\right) = i$ or $\pi\left(j\right) = i+1$. So, $|S_i| \le 6$, the important thing being that the size of $|S_i|$ is uniformly bounded. Moreover, $|T_i| \le C/\sqrt{n}$ for some constant C, which only depends on a and b. Next we examine

$$\sigma_n^2 = \text{Var}\left(\sum_{i=1}^n T_i\right) = \frac{1}{n}\text{Var}\left(\sum_{i=1}^n a\left[X_iX_{i+1} - 2pX_i\right] + b\left[X_{\pi(i)}X_{\pi(i+1)} - 2pX_{\pi(i)}\right]\right).$$

Again, these terms are for a given permutation π, though we don't explicitly include the conditioning operation in all variance and covariance expressions. By our previous calculations,

$$\sigma_n^2 = \left(a^2+b^2\right)p^2\left(1-p\right)^2 + o\left(1\right) + \frac{2ab}{n}\text{Cov}\left[\sum_{i=1}^{n-1} a\left(X_iX_{i+1} - 2pX_i\right), \sum_{i=1}^{n-1} b\left(X_{\pi(i)}X_{\pi(i+1)} - 2pX_{\pi(i)}\right)\right]. \quad \text{(C.7)}$$

We now show that the covariance term in (C.7) is $o\left(1\right)$. it can be expressed as

$$\frac{2ab}{n}\left[\text{Cov}\left(\sum_{i=1}^{n-1} X_iX_{i+1}, \sum_{j=1}^{n-1} X_{\pi(j)}X_{\pi(j+1)}\right) + 4p^2\,\text{Var}\left(\sum X_i\right) - 4p\,\text{Cov}\left(\sum_{i=1}^{n-1} X_iX_{i+1}, \sum_{j=1}^{n-1} X_j\right)\right]$$
$$= 2ab\left[\frac{1}{n}\text{Cov}\left(\sum_{i=1}^{n-1} X_iX_{i+1}, \sum_{j=1}^{n-1} X_{\pi(j)}X_{\pi(j+1)}\right) + 4p^3\left(1-p\right) - 2p^3\left(1-p\right) + o\left(1\right)\right]. \quad \text{(C.8)}$$

Next,

$$\frac{1}{n}\text{Cov}\left(\sum_{i=1}^{n-1} X_iX_{i+1}, \sum_{j=1}^{n-1} X_{\pi(j)}X_{\pi(j+1)}\right) = \frac{1}{n}\sum_{i=1}^{n-1}\sum_{j=1}^{n-1}\text{Cov}\left(X_iX_{i+1}, X_{\pi(j)}X_{\pi(j+1)}\right).$$

Now, for every i, there exists j such that $\pi\left(j\right) = i$. For such a j, if $\pi\left(j+1\right) \neq i+1$, then the

(i, j) covariance term in the double sum equals $p^3(1-p)$. Similarly, there exists a j such that $\pi(j+1) = i$. If $\pi(j) \neq i+1$, then the (i, j) term also equals $p^3(1-p)$. Similarly, there exists a j such that $\pi(j) = i+1$ and if $\pi(j+1) \neq i$, the term is also $p^3(1-p)$. Finally, there exists $j+1$ such that $\pi(j+1) = i+1$ and if $\pi(j) \neq i$, the term is also $p^3(1-p)$. So, the number of pairings of the above form is $4(n-1)$, except possibly if the permutation "preserves" ordering in the sense $\pi(j) = i$ and $\pi(j) = i+1$ or $\pi(j+1) = i$ and $\pi(j+1) = i+1$, in which case the covariance term would be $p^2 - p^4$. If, for a given sequence π_n, the number of such pairings is uniformly bounded above by some constant E, then (C.8) is equal to $4p^3(1-p) + o(1)$, and therefore (C.7) is equal to $o(1)$, as desired.

Unfortunately, we cannot simply argue that there is such a finite constant E, because if π is the identity permutation, the number of terms grows with n, though the identity permutation is an extremely unlikely outcome for a random permutation. Hence, we argue as follows. Let N_n be the number of indices (i, j) for which a random permutation Π satisfies $(\Pi(j), \Pi(j+1)) = (i, i+1)$ and similarly let N_n' be the number of $(\Pi(j+1), \Pi(j)) = (i+1, i)$. But,

$$\mathbb{E}[N_n] = \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} \mathbb{P}\{\Pi(j) = i, \Pi(j+1) = i = 1\} = \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} \frac{1}{n} \cdot \frac{1}{n-1} = \frac{n-1}{n},$$

and similarly for $\mathbb{E}'[N_n]$. Therefore, $M_n \equiv N_n + N_n'$ is uniformly bounded in expectation and, being nonnegative, is therefore tight. Hence, for any subsequence n_j, there exists a further subsequence for which M_n converges in distribution. By the almost sure representation theorem, there exists $\tilde{M}_n$ with the same distribution as M_n which converges almost surely to some $\tilde{M}$. Based on $\tilde{M} = m$, construct $\tilde{\Pi}_n$ according tot the conditional distribution of $\Pi_n | M_n = m$, so that unconditionally $\tilde{\Pi}_n$ is uniform over permutations. Then, along a subsequence, the above argument applies as if M were finite. To sum up, given any subsequence n_j there exists a further subsequence $n' = n_{jm}$ such that

$$\mathbb{P}\left\{\sum_{i=1}^{n} Z_i \leq t | \tilde{\Pi}_{n'}\right\} \to \Phi\left(t / \left[\sqrt{(a^2+b^2)} p(1-p)\right]\right)$$

with probability one. Hence, unconditionally, by dominated convergence,

$$\mathbb{P}\left\{\sum_{i=1}^{n} T_i \leq t\right\} \to \Phi\left(t / \left[\sqrt{(a^2+b^2)} p(1-p)\right]\right). \tag{C.9}$$

Since, given any subsequence, the same limit obtains for a further subsequence, the limit in (C.9) holds along the original subsequence. Hence, Hoeffding's condition holds.

Now for general k. That is, we consider the test statistic

$$\sqrt{n}\left[\hat{P}_{n,k}(\mathbf{X})-\left(1-\hat{Q}_{n,k}(\mathbf{X})\right)\right]=\sqrt{n}\left[\frac{\bar{V}_{n,k}}{\bar{V}_{n,k-1}}-\left(1-\frac{\bar{W}_{n,k}}{\bar{W}_{n,k-1}}\right)\right],$$

where

$$\bar{V}_{n,k}=\frac{1}{n}\sum_{i=1}^{n-k}X_i\cdots X_{i+k}$$

and

$$\bar{W}_{n,k}=\frac{1}{n}\sum_{i=1}^{n-k}(1-X_i)\cdots(1-X_{i+k}).$$

Using the Taylor approximation

$$\frac{\bar{V}_{n,k}}{\bar{V}_{n,k-1}}=p+p^{-k}\bar{V}_{n,k}-p^{1-k}\bar{V}_{n,k-1}+O_P\left(n^{-1}\right)$$

and similarly for $\bar{W}_{n,k}/\bar{W}_{n,k-1}$ yields

$$\hat{D}_{n,k}(\mathbf{X})=p^{-k}\bar{V}_{n,k}-p^{1-k}\bar{V}_{n,k-1}+(1-p)^{-k}\bar{W}_{n,k}-(1-p)^{1-k}\bar{W}_{n,k-1}+O_P\left(n^{-1}\right).$$

As in the proof of $k=1$, we need to verify Hoeffding's condition, i.e that $\sqrt{n}\hat{D}_{n,k}$ and $\sqrt{n}\tilde{D}_{n,k}$ are asymptotically independent, where

$$\tilde{D}^{\Pi}_{n,k}(\mathbf{X})=\hat{D}_{n,1}\left(X_{\Pi(1)},\ldots,X_{\Pi(n)}\right)$$

is the statistic $\hat{D}_{n,k}(\mathbf{X})$ evaluated at randomly permuted data $X_{\Pi(1)},\ldots,X_{\Pi(n)}$. As before, we first fix (or condition on) $\Pi=\pi$. For a given permutation π, define

$$\tilde{V}^{\pi}_{n,k}=\frac{1}{n}\sum_{i=1}^{n-k}X_{\pi(1)},\ldots,X_{\pi(i+k)},$$

and similarly for $\tilde{W}^{\pi}_{n,k}$. Using the same argument as in $k = 1$, we must show that

$$n\,\mathrm{Cov}\Big[p^{-k}\bar{V}_{n,k} - p^{1-k}\bar{V}_{n,k-1} + (1-p)^{-k}\,\bar{W}_{n,k} - (1-p)^{1-k}\,\bar{W}_{n,k-1}$$
$$p^{-k}\tilde{V}^{\pi}_{n,k} - p^{1-k}\tilde{V}^{\pi}_{n,k-1} + (1-p)^{-k}\,\tilde{W}^{\pi}_{n,k} - (1-p)^{1-k}\,\tilde{W}^{\pi}_{n,k-1}\Big] \to 0. \tag{C.10}$$

Before evaluating (C.10), which is a covariance between a linear combination of four random variables with another four random variables, we calculate the 16 individual covariances as follows:

$$n\,\mathrm{Cov}(\bar{V}_{n,k}, \tilde{V}^{\pi}_{n,k}) \to (k+1)^2 p^{2k+1}(1-p) \tag{C.11}$$
$$n\,\mathrm{Cov}(\bar{V}_{n,k}, \tilde{V}^{\pi}_{n,k-1}) \to k(k+1)p^{2k}(1-p)$$
$$n\,\mathrm{Cov}(\bar{V}_{n,k-1}, \tilde{V}^{\pi}_{n,k}) \to k(k+1)p^{2k}(1-p)$$
$$n\,\mathrm{Cov}(\bar{V}_{n,k-1}, \tilde{V}^{\pi}_{n,k-1}) \to k^2 p^{2k-1}(1-p)$$
$$n\,\mathrm{Cov}(\bar{W}_{n,k}, \tilde{W}^{\pi}_{n,k}) \to (k+1)^2 (1-p)^{2k+1} p$$
$$n\,\mathrm{Cov}(\bar{W}_{n,k}, \tilde{W}^{\pi}_{n,k-1}) \to k(k+1)(1-p)^{2k} p$$
$$n\,\mathrm{Cov}(\bar{W}_{n,k-1}, \tilde{W}^{\pi}_{n,k}) \to k(k+1)(1-p)^{2k} p$$
$$n\,\mathrm{Cov}(\bar{W}_{n,k-1}, \tilde{W}^{\pi}_{n,k-1}) \to k^2 (1-p)^{2k-1} p$$
$$n\,\mathrm{Cov}(\bar{V}_{n,k}, \tilde{W}^{\pi}_{n,k}) \to -(k+1)^2 p^{k+1}(1-p)^{k+1}$$
$$n\,\mathrm{Cov}(\bar{V}_{n,k}, \tilde{W}^{\pi}_{n,k-1}) \to -k(k+1)p^{k+1}(1-p)^{k}$$
$$n\,\mathrm{Cov}(\bar{V}_{n,k-1}, \tilde{W}^{\pi}_{n,k}) \to -k(k+1)p^{k}(1-p)^{k+1}$$
$$n\,\mathrm{Cov}(\bar{V}_{n,k-1}, \tilde{W}^{\pi}_{n,k-1}) \to -k^2 p^{k}(1-p)^{k}$$
$$n\,\mathrm{Cov}(\bar{W}_{n,k}, \tilde{V}^{\pi}_{n,k}) \to -(k+1)^2 p^{k+1}(1-p)^{k+1}$$
$$n\,\mathrm{Cov}(\bar{W}_{n,k-1}, \tilde{V}^{\pi}_{n,k}) \to -k(k+1)p^{k+1}(1-p)^{k}$$
$$n\,\mathrm{Cov}(\bar{W}_{n,k}, \tilde{V}^{\pi}_{n,k-1}) \to -k(k+1)p^{k}(1-p)^{k+1}$$
$$n\,\mathrm{Cov}(\bar{W}_{n,k-1}, \tilde{V}^{\pi}_{n,k-1}) \to -k^2 p^{k}(1-p)^{k}$$

The above 16 calculations are all similar, so we explain just (C.11). Note that

$$n\,\mathrm{Cov}(\bar{V}_{n,k}, \tilde{V}^{\pi}_{n,k}) = \frac{1}{n}\sum_{i=1}^{n-k}\sum_{j=1}^{n-k}\mathrm{Cov}\left(X_1\cdots X_{i+k}, X_{\pi(j)}\cdots X_{\pi(j+k)}\right). \tag{C.12}$$

Clearly, the (i, j) term in the double sum is zero if there is no overlap between the sets of indices

$\{i, \ldots, i+k\}$ and $\{\pi(j), \ldots, \pi(j+k)\}$. Then, for any fixed i, there exists some j such that $\pi(j) = i$, in which case there is overlap. Similarly, for any fixed l and m, each in $\{1, \ldots, k\}$ there exists j such that $\pi(j+m) = i+l$. Hence, there are $(k+1)^2$ sets of indices where there is some overlap between $\{i, \ldots, i+k\}$ and $\{\pi(j), \ldots, \pi(j+k)\}$. If for each combination of l and m, there is only one index that is shared, then the covariance is $p^{2k+1}(1-p)$. There are $(k+1)^2$ such terms. The only concern is there may be some further overlap in the sense that more than one index is shared among the sets. However, we will argue that, while this can occur, it has a small probability and hence will have negligible effect. To do this, for a given permutation π, let $N_n = N_n^\pi$ be the number of terms with at least two indices in common; that is, let

$$N_n^\pi = \sum_{i=1}^{n-k} \sum_{j=1}^{n-k} I\left\{ | \{i, \ldots, i+k\} \bigcap \{\pi(j), \ldots, \pi(j+k)\} | \geq 2 \right\}.$$

Since N_n^π can be bounded by the number of times pairs of indices are in common, we have

$$N_n^\pi \leq \sum_{i=1}^{n-k} \sum_{j=1}^{n-k} \sum_{l=0}^{k} \sum_{l'=0}^{k} \sum_{m=0}^{k} \sum_{m'=0}^{k} I\left\{ \pi(j+m) = i+l \bigcap \pi(j+m') = i+l' \right\}.$$

Then, continuing from (C.12),

$$|n \operatorname{Cov}(\bar{V}_{n,k}, \tilde{V}_{n,k}^\pi) - \frac{1}{n}(n-k)(k+1)^2 p^{2k+1}(1-p)| \leq N_n^\pi / n.$$

Therefore, for any given π(sequence), as long as $N_n^\pi/n \to 0$, the above 16 convergences hold.

Moreover, by the bilinearity of covariance, the above 16 convergences then allow us to calculate the LHS of (C.10) as the limit of the following 16 terms

$$\begin{aligned}
&np^{-2k} \operatorname{Cov}(\bar{V}_{n,k}, \tilde{V}_{n,k}^\pi) - np^{1-2k} \operatorname{Cov}(\bar{V}_{n,k}, \tilde{V}_{n,k-1}^\pi) + \cdots + n(1-p)^{2-2k} \operatorname{Cov}(\bar{W}_{n,k-1}, \tilde{V}_{n,k-1}^\pi) \\
&\to p(1-p)[(k+1)^2 - k(k+1) - (k+1)^2 + k(k+1) - k(k+1) + k^2 + k(k+1) - k^2 \\
&- (k+1)^2 + k(k+1) + (k+1)^2 - k(k+1) + k(k+1) - k^2 - k(k+1) + k^2] = 0
\end{aligned}$$

yielding the desired conclusion.

Thus, the covariance calculation implying the required asymptotic asymptotic independence in Hoeffding's condition holds, as long as $N_n^\pi/n \to 0$. (The asymptotic normality holds by the Central Limit Theorem of Stein 1986, used in the case $k = 1$ above.) This may not hold for every π (such as the identity permutation), but we now argue that, viewing N_n^Π as a function of

Π (i.e a random variable), $N_n^{\Pi}/n \to 0$ in probability. To see why,

$$\mathbb{E}\left[N_n^{\Pi}/n\right] \le \sum_{i=1}^{n-k}\sum_{j=1}^{n-k} \frac{(k+1)^2 k^2}{n(n+1)} \le (k+1)^2 k^2,$$

which is bounded in n. Thus, by Markov's inequality,

$$N_n^{\Pi}/n \xrightarrow{P} 0.$$

This does not guarantee $N_n^{\Pi}/n \to 0$ for almost every realization of the random permutation sequence Π (really Π_n), but we can use a subsequence argument as follows. Given any subsequence n_j, there exists a further subsequence $n' = n_{jm}$ such that $N_{n'} \to 0$ with probability one along the subsubsequence. Therefore, along the subsubsequence, we have that Hoeffding's condition holds for almost all realizations of Π'_n. We may then conclude, for any $\epsilon > 0$

$$\mathbb{P}\left\{|\hat{R}_{n'}(t) - \Phi(t/\sigma_k(p))| > \epsilon|\Pi_{n'}\right\} \to 0,$$

and so by dominated convergence

$$\hat{R}_{n'}(t) \to \Phi(t/\sigma_k(p))$$

with probability one along the subsubsequence. Since we can start with any subsequence before passing to a subsubsequence, and the limit remains the same, we can conclude that

$$\hat{R}_n(t) \xrightarrow{P} \Phi(t/\sigma_k(p)).$$

C.5 Proof of Theorem 2.2 (ii)

The proof is analogous to the proof of Theorem 2.2 (i). We provide the most important difference, the covariance calculation. Similar to $\hat{D}_{n,k}(\mathbf{X})$,

$$\hat{E}_{n,k}(\mathbf{X}) = \hat{P}_{n,k}(X_1, \ldots, X_n) - \hat{p}_n$$

admits a Taylor approximation given by

$$\hat{E}_{n,k}(\mathbf{X}) = p^{-k}\bar{V}_{n,k} - p^{k-1}\bar{V}_{n,k-1} - \bar{X}_n + O_P\left(n^{-1}\right).$$

Therefore, we need to show that, for fixed sequences $\pi = \pi_n$ (such that $N_n^\pi / n \to 0$ where N_n^π is defined as in the proof of Theorem 2.2 (i)),

$$n \operatorname{Cov}\left(p^{-k}\bar{V}_{n,k} - p^{k-1}\bar{V}_{n,k-1} - \bar{X}_n, p^{-k}\tilde{V}_{n,k}^\pi - p^{k-1}\tilde{V}_{n,k-1}^\pi - \tilde{X}_n^\pi\right) \to 0.$$

But the left side is

$$\begin{aligned} &2p^{1-2k}n \operatorname{Cov}(\bar{V}_{n,k}, \tilde{V}_{n,k-1}^\pi) - 2p^{-k}n \operatorname{Cov}(\bar{V}_{n,k}, \bar{X}_n) + 2p^{1-k}n \operatorname{Cov}(\tilde{V}_{n,k-1}^\pi, \bar{X}_n) \\ &+ nVar(\bar{X}_n) + p^{-2k}n \operatorname{Cov}(\bar{V}_{n,k}, \tilde{V}_{n,k}^\pi) + p^{2-2k}n \operatorname{Cov}(\bar{V}_{n,k-1}, \tilde{V}_{n,k-1}^\pi). \end{aligned} \tag{C.13}$$

Similar to the argument for (C.12), we have

$$n \operatorname{Cov}\left(\bar{X}_n, \bar{V}_{n,k}\right) \to (k+1)\, p^{k+1}\, (1-p)$$

and

$$n \operatorname{Cov}\left(\bar{X}_n, \bar{V}_{n,k-1}\right) \to k^2 p^{2k-1}\, (1-p)\, .$$

Plugging the limiting covariances into (C.13) yields the limiting variance of

$$\begin{aligned} &- 2p^{1-2k}k(k+1)p^{2k}(1-p) - 2p^{-k}(k+1)p^{k+1}(1-p) + 2p^{1-2k}kp^k(1-p) \\ &+ p(1-p) + p^{-2k}(k+1)^2p^{2k+1}(1-p) + p^{2-2k}k^2p^{2k-1}(1-p), \end{aligned}$$

which equals 0.

C.6 Proof of Theorem 3.1

As before, let $\mathbf{Y}_k = \left(Y_{ik} = \prod_{m=i}^{i+k} X_m, i \in \mathbb{Z}_+\right)$ and $\mathbf{Z}_k = \left(Z_{ik} = \prod_{m=i}^{i+k} (1 - X_m)\, , i \in \mathbb{Z}_+\right)$. By Theorem 5.2 of Bradley (2005),

$$\alpha\left(\mathbf{Y}_k, n\right) \le \sum_{m=0}^{k} \alpha\left(\mathbf{X}, n\right) \quad \text{and} \quad \alpha\left(\mathbf{Z}_k, n\right) \le \sum_{m=0}^{k} \alpha\left(\mathbf{X}, n\right)$$

Therefore, as $\mathbf{X}$ is α-mixing, $\alpha\left(\mathbf{Y}_k, n\right) \to 0$ and $\alpha\left(\mathbf{Z}_k, n\right) \to 0$ as $n \to \infty$ and $\mathbf{Y}_k$ and $\mathbf{Z}_k$ are α-mixing. Note also that $0 \le \alpha\left(\mathbf{G}, n\right) \le \frac{1}{4}$ for any sequence $\mathbf{G}$ and n. Thus, as $f(\delta) = x^{\frac{\delta}{2+\delta}}$ is decreasing in δ for $0 < x < 1$, for any $\delta > 0$,

$$\sum_{i=1}^{\infty} \left(\alpha\left(\mathbf{Y}_k, i\right)\right)^{\frac{\delta}{2+\delta}} \leq (k+1)^{\frac{\delta}{2+\delta}} \sum_{i=1}^{\infty} \left(\alpha\left(\mathbf{X}, i\right)\right)^{\frac{\delta}{2+\delta}} \leq (k+1)^{\frac{\delta}{2+\delta}} \sum_{i=1}^{\infty} \left(\alpha\left(\mathbf{X}, i\right)\right) < \infty,$$

and

$$\sum_{i=1}^{\infty} \left(\alpha\left(\mathbf{Z}_k, i\right)\right)^{\frac{\delta}{2+\delta}} \leq (k+1)^{\frac{\delta}{2+\delta}} \sum_{i=1}^{\infty} \left(\alpha\left(\mathbf{X}, i\right)\right)^{\frac{\delta}{2+\delta}} \leq (k+1)^{\frac{\delta}{2+\delta}} \sum_{i=1}^{\infty} \left(\alpha\left(\mathbf{X}, i\right)\right) < \infty.$$

First, we evaluate the asymptotic normal limiting distribution of $\hat{P}_{n,k}\left(\mathbf{X}\right) - \hat{p}$. Recall that $\mathit{\Gamma}_i = \left[Y_{ik}, Y_{i(k-1)}, X_i\right]^\top$. By Theorem 1.7 of Ibragimov (1962) and the Cramér-Wold device, as $\mathbf{Y}_k$ is strictly stationary with finite absolute moments and there exists a $\delta > 0$ such that $\sum_{i=1}^{\infty} \left(\alpha\left(\mathbf{Y}_k, i\right)\right)^{\frac{\delta}{2+\delta}} < \infty$,

$$n^{-1/2} \sum_{i=1}^{n} \mathit{\Gamma}_i \xrightarrow{d} N\left(\mathbb{E}\left[\mathit{\Gamma}_i\right], \Sigma\left(\mathit{\Gamma}_i\right)\right)$$

and each component of $\Sigma\left(\mathit{\Gamma}_i\right)$ is non-zero and finite. Therefore, by the Delta Method,

$$n^{1/2}\left(\hat{P}_{n,k}\left(\mathbf{X}\right) - \hat{p}\right) \xrightarrow{d} N\left(\frac{\mathbb{E}\left[Y_{ik}\right]}{\mathbb{E}\left[Y_{i(k-1)}\right]} - p, \mathbb{E}\left[\mathit{\Gamma}_i\right]^\top \Sigma\left(\mathit{\Gamma}_i\right) \mathbb{E}\left[\mathit{\Gamma}_i\right]\right).$$

Likewise, we evaluate the asymptotic normal limiting distribution of $\hat{D}_{n,k}\left(\mathbf{X}\right)$. Recall that $\mathit{\Lambda}_i = \left[Y_{ik}, Y_{i(k-1)}, Z_{ik}, Z_{ik-1}\right]^\top$. Again, by Theorem 1.7 of Ibragimov (1962) and the Cramér-Wold device, as $\mathbf{Z}_k$ is strictly stationary with finite absolute moments and there exists a $\delta > 0$ such that $\sum_{i=1}^{\infty} \left(\alpha\left(\mathbf{Z}_k, i\right)\right)^{\frac{\delta}{2+\delta}} < \infty$,

$$n^{-1/2} \sum_{i=1}^{n} \mathit{\Lambda}_i \xrightarrow{d} N\left(\mathbb{E}\left[\mathit{\Lambda}_i\right], \Sigma\left(\mathit{\Lambda}_i\right)\right),$$

where each component of $\Sigma\left(\mathit{\Lambda}_i\right)$ is non-zero and finite. Therefore, by the Delta Method,

$$n^{1/2}\left(\hat{D}_{n,k}\left(\mathbf{X}\right)\right) \xrightarrow{d} N\left(\frac{\mathbb{E}\left[Y_{ik}\right]}{\mathbb{E}\left[Y_{i(k-1)}\right]} - \left(1 - \frac{\mathbb{E}\left[Z_{ik}\right]}{\mathbb{E}\left[Z_{i(k-1)}\right]}\right), \mathbb{E}\left[\mathit{\Lambda}_i\right]^\top \Sigma\left(\mathit{\Lambda}_i\right) \mathbb{E}\left[\mathit{\Lambda}_i\right]\right).$$

C.7 Proof of Theorem 3.2

Note as well that the n-step transition matrix admits a closed form, given by

$$\mathcal{P}^n = \begin{bmatrix} \frac{1}{2} + \frac{1}{2}(2\epsilon)^n & \frac{1}{2} - \frac{1}{2}(2\epsilon)^n \\ \frac{1}{2} - \frac{1}{2}(2\epsilon)^n & \frac{1}{2} + \frac{1}{2}(2\epsilon)^n \end{bmatrix},$$

and that the stationarity of $\{X_i\}_{i=1}^n$ implies that X_1 is 0 or 1 with probabilities $1/2$.

We apply the delta method, and so need to compute the asymptotic variances and covariances of $\bar{V}_1$ and $\bar{V}_0$. First, we evaluate

$$\sqrt{n}\,\mathrm{Var}\left(\bar{X}_n\right) = \mathrm{Var}\left(X_1\right) + 2\sum_{i=1}^{n-1}\left(1 - \frac{i}{n}\right)\mathrm{Cov}\left(X_1, X_{i+1}\right).$$

For all $i \geq 1$,

$$\begin{aligned}
\mathbb{E}\left[X_1 X_{i+1}\right] &= \mathbb{E}\left[X_1 X_{i+1} | X = 1\right] + \mathbb{E}\left[X_1 X_{i+1} | X = 0\right] \\
&= \frac{1}{2}\mathbb{P}\left(X_{i+1} = 1 | X_1 = 1\right) \\
&= \frac{1}{2}\left(\frac{1}{2} + \frac{1}{2}(2\epsilon)^i\right)
\end{aligned}$$

which implies that

$$\mathrm{Cov}\left(X_1, X_{i+1}\right) = \frac{1}{2}\left(\frac{1}{2} + \frac{1}{2}(2\epsilon)^i\right) - \frac{1}{4} = \frac{1}{4}(2\epsilon)^i$$

and that therefore

$$\begin{aligned}
\sqrt{n}\,\mathrm{Var}\left(\bar{X}_n\right) &= \frac{1}{4} + 2\sum_{i=1}^{n-1}\left(1 - \frac{i}{n}\right)\frac{1}{4}(2\epsilon)^i \\
&\to \frac{1}{4} + \frac{1}{2}\sum_{i=1}^{\infty}(2\epsilon)^i \\
&= \frac{1}{4} + \frac{1}{2}\left[\frac{1}{1-2\epsilon} - 1\right] \\
&= \frac{1}{4} + \frac{\epsilon}{1-2\epsilon}.
\end{aligned}$$

Next, we evaluate

$$\sqrt{n}\,\mathrm{Var}\left(\sum_{i=1}^{n-1} X_i X_{i+1}\right) = \mathrm{Var}\left(X_1 X_2\right) + 2\sum_{i=1}^{n-1}\left(1-\frac{i}{n-1}\right)\mathrm{Cov}\left(X_1X_2, X_{i+1}X_{i+2}\right).$$

In order to do so, we would need to evaluate $\mathbb{E}\left[X_1X_2X_{i+1}X_{i+2}\right]$. If we set $i=1$, then

$$\begin{aligned}
\mathbb{E}\left[X_1X_2X_{i+1}X_{i+2}\right] &= \mathbb{E}\left[X_1X_2X_3\right] \\
&= \frac{1}{2}\mathbb{E}\left[X_1X_2X_3 | X_1 = 1\right] \\
&= \frac{1}{2}\left(\frac{1}{2}+\epsilon\right)^2,
\end{aligned}$$

and we can evaluate

$$\begin{aligned}
\mathrm{Cov}\left(X_1X_2, X_2X_3\right) &= \frac{1}{2}\left(\frac{1}{2}+\epsilon\right)^2 - \frac{1}{4}\left(\frac{1}{2}+\epsilon\right)^2 \\
&= \frac{1}{2}\left(\frac{1}{2}+\epsilon\right).^2
\end{aligned}$$

If we set $i>1$, then

$$\mathbb{E}\left[X_1X_2X_{i+1}X_{i+2}\right] = \frac{1}{2}\left(\frac{1}{2}+\frac{1}{2}\left(2\epsilon\right)^{i-1}\right)\left(\frac{1}{2}+\epsilon\right)^2$$

and

$$\begin{aligned}
\mathrm{Cov}\left(X_1X_2, X_{i+1}X_{i+2}\right) &= \frac{1}{2}\left(\frac{1}{2}+\frac{1}{2}\left(2\epsilon\right)^{i-1}\right)\left(\frac{1}{2}+\epsilon\right)^2 - \frac{1}{4}\left(\frac{1}{2}+\epsilon\right)^2 \\
&= \frac{1}{4}\left(2\epsilon\right)^{i-1}\left(\frac{1}{2}+\epsilon\right)^2.
\end{aligned}$$

Therefore, we see that

$$\begin{aligned}
\sqrt{n}\,\mathrm{Var}\left(\sum_{i=1}^{n-1} X_iX_{i+1}\right) &= \frac{1}{2}\left(\frac{1}{2}+\epsilon\right)\left(1-\frac{1}{2}\left(\frac{1}{2}+\epsilon\right)\right) + \frac{1}{2}\sum_{i=1}^{n-2}\left(1-\frac{i}{n-1}\right)\left(2\epsilon\right)^{i-1}\left(\frac{1}{2}+\epsilon\right)^2 \\
&\to \left(\frac{1}{4}+\frac{\epsilon}{2}\right)\left(\frac{3}{4}-\frac{\epsilon}{2}\right) + \frac{1}{2}\left(\frac{1}{2}+\epsilon\right)^2\sum_{i=0}^{\infty}\left(2\epsilon\right)^i \\
&= \left(\frac{1}{4}+\frac{\epsilon}{2}\right)\left(\frac{3}{4}-\frac{\epsilon}{2}\right) + \frac{1}{2}\left(\frac{1}{2}+\epsilon\right)^2\frac{1}{1-2\epsilon}.
\end{aligned}$$

Finally, we evaluate

$$\operatorname{Cov}\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{n-1} X_i X_{j+1}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n-1} \operatorname{Cov}\left(X_i, X_j X_{j+1}\right). \tag{C.14}$$

In order to do so, we need to evaluate $\mathbb{E}\left[X_i X_j X_{j+1}\right]$. If we set $i = j$, then $\mathbb{E}\left[X_j X_{j+1}\right] = \frac{1}{2}\left(\frac{1}{2}+\epsilon\right)$ and $\operatorname{Cov}\left(X_j, X_j X_{j+1}\right) = \frac{1}{2}\left(\frac{1}{2}+\epsilon\right) - \frac{1}{4}\left(\frac{1}{2}-\epsilon\right) = \frac{1}{4}\left(\frac{1}{2}+\epsilon\right)$. Similarly, if $i = j+1$, then $\mathbb{E}\left[X_j X_{j+1}\right] = \frac{1}{2}\left(\frac{1}{2}+\epsilon\right)$ and $\operatorname{Cov}\left(X_{j+1}, X_j X_{j+1}\right) = \frac{1}{4}\left(\frac{1}{2}+\epsilon\right)$. If we set $i < j$, then

$$\begin{aligned}\mathbb{E}\left[X_i X_j X_{j+1}\right] &= \frac{1}{2}\left(\frac{1}{2} + \frac{(2\epsilon)^{j-i}}{2}\right)\left(\frac{1}{2}+\epsilon\right)\\ &= \frac{1}{4}\left(1 + (2\epsilon)^{j-i}\right)\left(\frac{1}{2}+\epsilon\right)\end{aligned}$$

and

$$\begin{aligned}\operatorname{Cov}\left(X_i X_j X_{j+1}\right) &= \frac{1}{4}\left(1 + (2\epsilon)^{j-i}\right)\left(\frac{1}{2}+\epsilon\right) - \frac{1}{4}\left(\frac{1}{2}+\epsilon\right)\\ &= \frac{1}{4}\left((2\epsilon)^{j-i}\right)\left(\frac{1}{2}+\epsilon\right).\end{aligned}$$

If we set $i > j+1$, then

$$\begin{aligned}\mathbb{E}\left[X_i X_j X_{j+1}\right] &= \frac{1}{2}\left(\frac{1}{2} + \frac{(2\epsilon)^{i-j-1}}{2}\right)\left(\frac{1}{2}+\epsilon\right)\\ &= \frac{1}{4}\left(1 + (2\epsilon)^{i-j-1}\right)\left(\frac{1}{2}+\epsilon\right)\end{aligned}$$

and

$$\begin{aligned}\operatorname{Cov}\left(X_i X_j X_{j+1}\right) &= \frac{1}{4}\left(1 + (2\epsilon)^{i-j-1}\right)\left(\frac{1}{2}+\epsilon\right) - \frac{1}{4}\left(\frac{1}{2}+\epsilon\right)\\ &= \frac{1}{4}\left((2\epsilon)^{i-j-1}\right)\left(\frac{1}{2}+\epsilon\right).\end{aligned}$$

Therefore, we have shown

$$\operatorname{Cov}\left(X_i X_j X_{j+1}\right) = \begin{cases}\frac{1}{4}\left((2\epsilon)^{j-i}\right)\left(\frac{1}{2}+\epsilon\right) & \text{if } i \leq j\\ \frac{1}{4}\left((2\epsilon)^{i-j-1}\right)\left(\frac{1}{2}+\epsilon\right) & \text{if } i > j.\end{cases}$$

We can evaluate (C.14) by splitting the summation.

$$\sum_{i=1}^{n}\sum_{j=1}^{n-1}\operatorname{Cov}\left(X_i, X_jX_{j+1}\right) = \sum_{i=1}^{n}\sum_{j=i}^{n-1}\operatorname{Cov}\left(X_i, X_jX_{j+1}\right) + \sum_{i=2}^{n}\sum_{j=1}^{i-1}\operatorname{Cov}\left(X_i, X_jX_{j+1}\right). \tag{C.15}$$

The first term of (C.15) can be evaluated at the limit as

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n}\sum_{j=i}^{n-1}\operatorname{Cov}\left(X_i, X_jX_{j+1}\right) &= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n-1}\frac{1}{4}\left(2\epsilon\right)^{j-i}\left(\frac{1}{2}+\epsilon\right)\\
&= \frac{1}{n}\sum_{i=1}^{n}\sum_{k=0}^{n-i-1}\frac{1}{4}\left(2\epsilon\right)^{k}\left(\frac{1}{2}+\epsilon\right)\\
&= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{4}\left(\frac{1}{2}+\epsilon\right)\sum_{k=0}^{n-i-1}\left(2\epsilon\right)^{k}\\
&= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{4}\left(\frac{1}{2}+\epsilon\right)\frac{1-\left(2\epsilon\right)^{n-i}}{1-\left(2\epsilon\right)}\\
&\to \frac{1}{4}\left(\frac{1}{2}+\epsilon\right)\frac{1}{1-2\epsilon}.
\end{aligned}$$

Likewise, the second term of (C.15) can be evaluated at the limit as

$$\begin{aligned}
\frac{1}{n}\sum_{i=2}^{n}\sum_{j=1}^{i-1}\operatorname{Cov}\left(X_i, X_jX_{j+1}\right) &= \frac{1}{n}\sum_{i=2}^{n}\sum_{j=1}^{i-1}\frac{1}{4}\left(2\epsilon\right)^{i-j-1}\left(\frac{1}{2}+\epsilon\right)\\
&= \frac{1}{n}\sum_{i=2}^{n}\frac{1}{4}\left(\frac{1}{2}+\epsilon\right)\sum_{k=1}^{i-1}\left(2\epsilon\right)^{k-1}\\
&= \frac{1}{n}\sum_{i=1}^{n}\frac{1}{4}\left(\frac{1}{2}+\epsilon\right)\frac{1-\left(2\epsilon\right)^{i-1}}{1-\left(2\epsilon\right)}\\
&\to \frac{1}{4}\left(\frac{1}{2}+\epsilon\right)\frac{1}{1-2\epsilon}.
\end{aligned}$$

Hence, we can sum the limits to see that

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n-1}\operatorname{Cov}\left(X_i, X_jX_{j+1}\right) \to \frac{1}{2}\left(\frac{1}{2}+\epsilon\right)\frac{1}{1-2\epsilon}.$$

As $\{X_i\}_{i=1}^{n}$ and $\{X_iX_{i+1}\}_{i=1}^{n-1}$ are irreducible and aperiodic, they are both α-mixing by Theorem 3.1 of Bradley (2005). Therefore, by Theorem B.0.1 of Politis et. al (1999), a central limit theorem for $\alpha-$mixing triangular arrays, and the Cramér-Wold device, we have that

$$\sqrt{n}\left(\bar{X}_n-\frac{1}{2},\frac{\sum_{i=1}^{n-1}X_iX_{i+1}}{n-1}-\frac{1}{2}\left(\frac{1}{2}-\epsilon\right)\right)\xrightarrow{d}N\left(0,\Sigma\right)$$

where

$$\Sigma=\begin{bmatrix}\frac{1}{4}+\frac{\epsilon}{1-2\epsilon} & \frac{1}{2}\left(\frac{1}{2}+\epsilon\right)\frac{1}{1-2\epsilon}\\ \frac{1}{2}\left(\frac{1}{2}+\epsilon\right)\frac{1}{1-2\epsilon} & \left(\frac{1}{4}+\frac{\epsilon}{2}\right)\left(\frac{3}{4}-\frac{\epsilon}{2}\right)+\frac{1}{2}\left(\frac{1}{2}+\epsilon\right)^2\frac{1}{1-2\epsilon}\end{bmatrix}.$$

Next, we apply the delta method to evaluate the limiting distribution of $\hat{D}_{n,1}\left(\mathbf{X}\right)$. Note that, as in the proof of Theorem 2.2 (i),

$$\sqrt{n}\left(\hat{D}_{n,1}\left(\mathbf{X}\right)-2\epsilon\right)=\sqrt{n}\left(\frac{\bar{V}_1-\bar{V}_0^2}{\bar{V}_0\left(1-\bar{V}_0\right)}-2\epsilon\right)+o_p(1).$$

Let

$$f(v_0,v_1)=\frac{v_1-v_0^2}{v_0\left(1-v_0\right)}$$

and define $\mu_1=\mathbb{E}\left[\bar{V}_1\right]=\frac{1}{2}\left(\frac{1}{2}+\epsilon\right)$ and $\mu_0=\mathbb{E}\left[\bar{V}_0\right]=\frac{1}{2}$. We can evaluate

$$f\left(\mu_0,\mu_1\right)=\frac{\frac{1}{2}\left(\frac{1}{2}+\epsilon\right)-\frac{1}{4}}{1/4}=2\epsilon,$$

$$\frac{\partial f}{\partial v_0}\Big|_{\mu}=\frac{-2v_0^2\left(1-v_0\right)-\left(v_1-v_0^2\right)\left(1-2v_0\right)}{v_0^2\left(1-v_0\right)^2}\Big|_{\mu}=-4,$$

and

$$\frac{\partial f}{\partial v_1}\Big|_{\mu}=\frac{1}{v_0\left(1-v_0\right)}\Big|_{\mu}=4.$$

Therefore, we can see that

$$\sqrt{n}\left(\hat{D}_{n,1}\left(\mathbf{X}\right)-2\epsilon\right)=\sqrt{n}\left(4\left(\bar{V}_1-\mu_1\right)-4\left(\bar{V}_0-\mu_0\right)\right)+o_p\left(1\right).$$

Note that

$$\begin{aligned}&\mathrm{Var}\left(4\left(\bar{V}_1-\mu_1\right)-4\left(\bar{V}_0-\mu_0\right)\right)\\&=16\left(\frac{1}{4}+\frac{\epsilon}{1-2\epsilon}+\left(\frac{1}{4}+\frac{\epsilon}{2}\right)\left(\frac{3}{4}-\frac{\epsilon}{2}\right)+\frac{1}{2}\left(\frac{1}{2}+\epsilon\right)^2\frac{1}{1-2\epsilon}-2\left(\frac{1}{2}\right)\left(\frac{1}{2}+\epsilon\right)\frac{1}{1-2\epsilon}\right)\\&=1-4\epsilon^2.\end{aligned}$$

Hence, we have that

$$n^{1/2}\left(\hat{D}_{n,1}\left(\mathbf{X}\right)-2\epsilon\right)\xrightarrow{d}N\left(0,1-4\epsilon^{2}\right).$$

Again, we apply the delta method to evaluate the limiting distribution of $\hat{P}_{n,1}$. Let

$$g(v_0,v_1)=\frac{v_1}{v_0}.$$

We can evaluate

$$g\left(\mu_0,\mu_1\right)=\frac{1}{2}+\epsilon,$$

$$\frac{\partial g}{\partial v_0}\Big|_{\mu}=\frac{-v_1}{v_0^2}\Big|_{\mu}=-1-2\epsilon,$$

and

$$\frac{\partial g}{\partial v_1}\Big|_{\mu}=\frac{1}{v_0}\Big|_{\mu}=2.$$

Therefore, we can see that

$$\sqrt{n}\left(\hat{P}_{n,1}-\frac{1}{2}-\epsilon\right)=\sqrt{n}\left(2\left(\bar{V}_1-\mu_1\right)-\left(1+2\epsilon\right)\left(\bar{V}_0-\mu_0\right)\right)+o_p\left(1\right).$$

Note that

$$\begin{aligned}
&\operatorname{Var}\left(2\left(\bar{V}_1-\mu_1\right)-\left(1+2\epsilon\right)\left(\bar{V}_0-\mu_0\right)\right)\\
&=4\left(\left(\frac{1}{4}+\frac{\epsilon}{2}\right)\left(\frac{3}{4}-\frac{\epsilon}{2}\right)+\frac{1}{2}\left(\frac{1}{2}+\epsilon\right)^2\frac{1}{1-2\epsilon}\right)\\
&+\left(1+2\epsilon\right)^2\left(\frac{1}{4}+\frac{\epsilon}{1-2\epsilon}\right)-4\left(1+2\epsilon\right)\left(\frac{1}{2}\left(\frac{1}{2}+\epsilon\right)\frac{1}{1-2\epsilon}\right)\\
&=\frac{1}{2}-2\epsilon^2.
\end{aligned}$$

Hence, we have that

$$n^{1/2}\left(\hat{P}_{n,1}\left(\mathbf{X}\right)-\frac{1}{2}-\epsilon\right)\xrightarrow{d}N\left(0,\frac{1}{2}-2\epsilon^{2}\right).$$

Finally, we apply the delta method to evaluate the limiting distribution of $\hat{P}_{n,1}-\bar{X}_n$. Let

$$h(v_0,v_1)=\frac{v_1}{v_0}-v_0$$

We can evaluate

$$h\left(\mu_0, \mu_1\right) = \epsilon,$$

$$\left.\frac{\partial h}{\partial v_0}\right|_{\mu} = \left.\frac{-v_1}{v_0^2} - 1\right|_{\mu} = -2 - 2\epsilon,$$

and

$$\left.\frac{\partial h}{\partial v_1}\right|_{\mu} = \left.\frac{1}{v_0}\right|_{\mu} = 2.$$

Therefore, we can see that

$$\sqrt{n}\left(\hat{P}_{n,1}\left(\mathbf{X}\right) - \hat{p}_n - \epsilon\right) = \sqrt{n}\left(2\left(\bar{V}_1 - \mu_1\right) - \left(2 + 2\epsilon\right)\left(\bar{V}_0 - \mu_0\right)\right) + o_p\left(1\right).$$

Note that

$$\begin{aligned}
&\mathrm{Var}\left(2\left(\bar{V}_1 - \mu_1\right) - \left(2 + 2\epsilon\right)\left(\bar{V}_0 - \mu_0\right)\right) \\
&= 4\left(\left(\frac{1}{4} + \frac{\epsilon}{2}\right)\left(\frac{3}{4} - \frac{\epsilon}{2}\right) + \frac{1}{2}\left(\frac{1}{2} + \epsilon\right)^2 \frac{1}{1 - 2\epsilon}\right) \\
&+ \left(2 + 2\epsilon\right)^2\left(\frac{1}{4} + \frac{\epsilon}{1 - 2\epsilon}\right) - 4\left(2 + 2\epsilon\right)\left(\frac{1}{2}\left(\frac{1}{2} + \epsilon\right)\frac{1}{1 - 2\epsilon}\right) \\
&= \frac{1 - 2\epsilon + 16\epsilon^2}{4 - 8\epsilon}.
\end{aligned}$$

Hence, we have that

$$n^{1/2}\left(\hat{P}_{n,1} - \bar{X}_n - \epsilon\right) \xrightarrow{d} N\left(0, \frac{1 - 2\epsilon + 16\epsilon^2}{4 - 8\epsilon}\right).$$

C.8 Proof of Theorem 3.3

We can rewrite the problem as a 2^m state Markov Chain of order 2 on $\prod_{i=1}^m \{0,1\}_i$. For example, suppose $m = 2$, then a transition from $\{0,1\}$ to $\{1,0\}$ occurs in the second position of the sequence $\{0,1,0\}\}$.

Let the j^{th} state be a m-tuple on $\{0,1\}$ denoted by $\left\{I_m^j, \ldots, I_1^j\right\}$. The states are enumerated such that $I_m^j I_{m-1}^j \ldots I_1^j$ is $j-1$ expressed in the base 2 numeral system, i.e. $\{1,0,0,0\}$ is the 9^{th} state when $m = 4$. Throughout this proof, we let j denote the state $\left\{I_m^j, \ldots, I_1^j\right\}$.

Let $l = 2^{m-1}$ and $j' = j \pmod{l}$. If the Markov Chain is in state j, then it can only transition to states $2j' - 1 = \left\{I_{m-1}^j, \ldots, 0\right\}$ and $2j' = \left\{I_{m-1}^j, \ldots, 1\right\}$, as appending a 0 to the end of a integer expressed base 2 is equivalent to multiplication by $2,$ removing the first digit of an integer $\geq 2^{m-1}$ and $\leq 2^m$ expressed base 2 is equivalent to taking $\pmod{l}$, and $\{I_m^{2r}, \ldots, I_2^{2r}\} = \left\{I_m^{2r-1}, \ldots, I_2^{2r-1}\right\}$ for all $0 \leq r \leq l$.

Let the stationary distribution of the Markov Chain be denoted by $\pi = (\pi_1, \ldots, \pi_{2^m})$ and the probability from transitioning from state j to state s be denoted by $\rho(j, s)$. In general, π must satisfy the system of $2^m + 1$ equations

$$\begin{aligned}
\pi_1 \rho(1,1) + \pi_{l+1} \rho(l+1,1) &= \pi_1 \\
\pi_1 \rho(1,2) + \pi_{l+1} \rho(l+1,2) &= \pi_1 \\
&\vdots \\
\pi_j \rho(j, 2j-1) + \pi_{l+j} \rho(l+j, 2j-1) &= \pi_{2j-1} \\
\pi_j \rho(j, 2j) + \pi_{l+j} \rho(l+j, 2j) &= \pi_{2j} \\
&\vdots \\
\pi_l \rho(l, 2^m-1) + \pi_{2^m} \rho(2^m, 2^m-1) &= \pi_{2^m-1} \\
\pi_l \rho(l, 2^m) + \pi_{2^m} \rho(2^m, 2^m) &= \pi_{2^m} \\
\sum_{j=1}^{2^m} \pi_j &= 1
\end{aligned} \tag{C.16}$$

In the Markov Chain we consider, $\rho(1,1) = \rho\left(2^k, 2^k\right) = \frac{1}{2} + \epsilon$, $\rho(1,2) = \rho\left(2^k, 2^k - 1\right) = \frac{1}{2} - \epsilon$, $\rho(j, 2j-1) = \rho(j, 2j) = \rho(l+j, 2j-1) = \rho(l+j, 2j) = \frac{1}{2}$ for all j in $1, \ldots l$, and all other transition probabilities be equal to 0.

We find that, in this model,

$$\pi_j = \begin{cases} \frac{1}{2+(2^m-2)(1-2\epsilon)} & \text{for } j = 1, 2^m \\ \frac{1-2\epsilon}{2+(2^m-2)(1-2\epsilon)} & \text{for } 2 \leq j \leq 2^{m-1}. \end{cases} \tag{C.17}$$

It is straightforward to show that the stationary distribution given by (C.17) satisfies the system of equations (C.16) and is therefore the stationary distribution of the Markov Chain under consideration.

Let $\mathcal{I}^m = \{0,1\}^m$ be the set of sequences of zeros and ones of length m. Let $\mathcal{I}_q^m \subset \mathcal{I}^m$ be the set of sequences of zeros and ones of length m where the first q elements are ones, the $q+1^{\text{th}}$ element is a 0, and elements $q+2$ through m are either zeros or ones. Note that $\{\mathcal{I}_0^m, \ldots, \mathcal{I}_k^m\}$

is a partition of $\mathcal{I}^m$ and that the cardinality of $\mathcal{I}_q^m$ is 2^{m-q-1} for $0 \leq q < k$ and 1 for $q = m$. Finally, let $\mathcal{I}_{1+}^m = \mathcal{I}^m \backslash \mathcal{I}_0^m$.

We can see that, for $k \geq m-1$,

$$\begin{aligned}
\mathbb{E}\left[Y_{jk}\right] &= \sum_{I \in \mathcal{I}_{1+}^m} \mathbb{P}\left(X_{j+1} = 1, \ldots, X_{j+k} = 1 \middle| \left(X_j, \ldots, X_{j-m+1}\right) = I\right) \mathbb{P}\left(\left(X_j, \ldots, X_{j-m+1}\right) = I\right) \\
&= \sum_{q=1}^{m-1} 2^{m-q-1} \left(\frac{1}{2}\right)^{m-q} \left(\frac{1}{2} + \epsilon\right)^{k-m+q} \left(\frac{1-2\epsilon}{2 + \left(2^m - 2\right)\left(1 - 2\epsilon\right)}\right) \\
&+ \left(\frac{1}{2} + \epsilon\right)^{k} \left(\frac{1}{2 + \left(2^m - 2\right)\left(1 - 2\epsilon\right)}\right) \\
&= \frac{\left(\frac{1}{2} + \epsilon\right)^{k-m+1}}{2 + \left(2^m - 2\right)\left(1 - 2\epsilon\right)}.
\end{aligned}$$

Likewise, for $1 \leq k < m-1$,

$$\begin{aligned}
\mathbb{E}\left[Y_{jk}\right] &= \sum_{I \in \mathcal{I}_{1+}^m} \mathbb{P}\left(X_{j+1} = 1, \ldots, X_{j+k} = 1 \middle| \left(X_j, \ldots, X_{j-m+1}\right) = I\right) \mathbb{P}\left(\left(X_j, \ldots, X_{j-m+1}\right) = I\right) \\
&= \sum_{q=1}^{m-k} 2^{m-q-1} \left(\frac{1}{2}\right)^{k} \left(\frac{1-2\epsilon}{2 + \left(2^m - 2\right)\left(1 - 2\epsilon\right)}\right) \\
&+ \sum_{q=m-k+1}^{m-1} 2^{m-q-1} \left(\frac{1}{2}\right)^{m-q} \left(\frac{1}{2} + \epsilon\right)^{k-m+q} \left(\frac{1-2\epsilon}{2 + \left(2^m - 2\right)\left(1 - 2\epsilon\right)}\right) \\
&+ \left(\frac{1}{2} + \epsilon\right)^{k} \left(\frac{1}{2 + \left(2^m - 2\right)\left(1 - 2\epsilon\right)}\right) \\
&= \frac{2^{m-k-1} + \left(2 - 2^{m-k}\right)\epsilon}{2 + \left(2^m - 2\right)\left(1 - 2\epsilon\right)}.
\end{aligned}$$

Finally, $\mathbb{E}\left[Y_{jk}\right] = \frac{1}{2}$ for $k = 0$.

Therefore, by Theorem 3.1,

$$n^{1/2}\left(\hat{P}_{n,k}\left(\mathbf{X}\right) - \mu_P\left(k, m, \epsilon\right)\right) \xrightarrow{d} N\left(0, \sigma_P^2\left(k, m, \epsilon\right)\right),$$

where

$$\mu_P(k,m,\epsilon) = \begin{cases} \epsilon + \frac{1}{2} & \text{if } m \leq k \\ \frac{1}{2(1-\epsilon)} & \text{if } k = m-1 \\ \frac{2^{m-k-1}+\left(2-2^{m-k}\right)\epsilon}{2^{m-k}+\left(2-2^{m-k+1}\right)\epsilon} & \text{if } 1 < k < m-1 \\ \frac{2^{m-1}+(4-2^m)\epsilon}{2+(2^m-2)(1-2\epsilon)} & \text{if } k = 1 \end{cases}$$

and $\sigma_P^2(k,m,\epsilon)$ is a function of k, m, and ϵ. $\hat{P}_{n,k}(\mathbf{X}) - \hat{p}$ is asymptotically normal with limiting distribution given by

$$n^{1/2}\left(\hat{P}_{n,k}(\mathbf{X}) - \hat{p} - \mu_{\tilde{P}}(k,m,\epsilon)\right) \xrightarrow{d} N\left(0, \sigma_{\tilde{P}}^2(k,m,\epsilon)\right),$$

where

$$\mu_{\tilde{P}}(k,m,\epsilon) = \begin{cases} \epsilon & \text{if } m \leq k \\ \frac{\epsilon}{2(1-\epsilon)} & \text{if } k = m-1 \\ \frac{\epsilon}{2^{m-k}+\left(2-2^{m-k+1}\right)\epsilon} & \text{if } 1 < k < m-1 \\ \frac{2}{2+(2^m-2)(1-2\epsilon)} & \text{if } k = 1 \end{cases}$$

and $\sigma_{\tilde{P}}^2(k,m,\epsilon)$ is a function of k, m, and ϵ. $\hat{D}_{n,k}(\mathbf{X})$ is asymptotically normal with limiting distribution given by

$$n^{1/2}\left(\hat{D}_{n,k}(\mathbf{X}) - \mu_D(k,m,\epsilon)\right) \xrightarrow{d} N\left(0, \sigma_D^2(k,m,\epsilon)\right).$$

where

$$\mu_D(k,m,\epsilon) = \begin{cases} 2\epsilon & \text{if } m \leq k \\ \frac{\epsilon}{1-\epsilon} & \text{if } k = m-1 \\ \frac{\epsilon}{2^{m-k-1}+\left(1-2^{m-k}\right)\epsilon} & \text{if } 1 < k < m-1 \\ \frac{4\epsilon}{2+(2^m-2)(1-2\epsilon)} & \text{if } k = 1 \end{cases}$$

and $\sigma_D^2(k,m,\epsilon)$ is a function of k, m, and ϵ.

C.9 Proof of Theorem 3.4

As before, let $V_{n,1} = \sum_{i=1}^{n-1} X_i X_{i+1}$ and $V_{n,0} = \sum_{i=1}^{n} X_i$. Let R_0 denote the number of runs of zeros and R_1 the number of runs of ones, so that $R = R_0 + R_1$. Since the first success in a run of ones does not contribute to the sum $V_{n,1}$, the number of ones followed by a one in a particular run of ones is the number of ones in the run minus one. Therefore, $V_{n,1} = V_{n,0} - R_1$. So, if R

is even, $V_{n,1} = V_{n,0} - R/2$. On the other hand, if R is odd, then there are either $(R-1)/2$ or $(R+1)/2$ runs of ones. It follows that

$$|V_{n,1} - \left(V_{n,0} - \frac{R}{2}\right)| \leq \frac{1}{2}$$

or

$$\frac{V_{n,1}}{n} = \hat{p} - \frac{R}{2n} + O_P\left(n^{-1}\right) \tag{C.18}$$

where as before $\hat{p} = V_{n,0}/n$. In order to show (3.8), by C.5, it suffices to show

$$\sqrt{n}\left[\left(\frac{\frac{V_{n,1}}{n} - \hat{p}^2}{\hat{p}\left(1-\hat{p}\right)}\right) + \frac{\frac{R}{2n} - \hat{p}\left(1-\hat{p}\right)}{\hat{p}\left(1-\hat{p}\right)}\right] \overset{P}{\to} 0,$$

or equivalently

$$\sqrt{n}\left[\frac{V_{n,1}}{n} - \hat{p}^2 + \frac{R}{2n} - \hat{p}\left(1-\hat{p}\right)\right] = \sqrt{n}\left[\frac{V_{n,1}}{n} + \frac{R}{2n} - \hat{p}\right] \overset{P}{\to} 0,$$

which now follows trivially from (C.18). Thus (i) holds and (ii) trivially follows. Part (iii) follows from Slutsky's theorem from randomization distributions; see Chung and Romano (2013), Theorem 5.2. Indeed, from Theorem 3.2 we know the permutation distribution based on $\sqrt{n}\hat{D}_{n,1}\left(\mathbf{X}\right)$ is asymptotically $N\left(0,1\right)$ (in probability) and the same must then be true for $\sqrt{n}Z_n$. Part (iv) then follows because the critical values are also asymptotically (or exactly) $z_{1-\alpha}$ under contiguous alternatives.

C.10 Proof of Theorem 3.5

In order to deduce the result, we need to establish an asymptotic equicontinuity property of $L_n\left(h\right)$. Let $\rho\left(F,G\right)$ denote the Levy metric between distributions F and G.

Lemma C.1. *Given any $\delta > 0$, there exists $\delta' > 0$ such that if $|h_n - h| \leq \delta'$, then*

$$\rho\left(L_n\left(h_n\right), L_n\left(h\right)\right) \leq \delta$$

for sufficiently large n.

We use the following coupling of data sets based on $S\left(h_n\right)$ number of successes and also $S\left(h\right)$ number of successes. That is, let $\mathbf{X} = \left(X_1, \ldots, X_n\right)$ be the data set with the first $S =$

$S(h_n)$ entries equal to 1 and the rest 0. Similarly, let $\mathbf{X}' = (X'_1, \ldots, X'_n)$ be the data set with the first $S' = S(h)$ entries equal to 1 and the rest 0. We first claim that, for any $\delta > 0$ and any permutation π applied to both X and X', there exists $\delta' > 0$ independent of π, such that for sufficiently large n independent of π,

$$\sqrt{n}|\hat{D}_{n,1}(\mathbf{X}_\pi) - \hat{D}_{n,1}(\mathbf{X}'_\pi)| \leq \delta \tag{C.19}$$

if $|h_n - h| \leq \delta'$. The lemma would then follow because $L_n(h)$ puts equal mass at the values $\sqrt{n}\hat{D}_{n,1}(\mathbf{X}_\pi)$ as πvaries while $L_n(h)$ puts equal mass at the values $\sqrt{n}\hat{D}_{n,1}(\mathbf{X}'_\pi)$. In general, if F puts mass $1/N$ at data points $a_1, \ldots, a_N$ and G puts mass $1/N$ at data points $b_1, \ldots, b_N$, with $|b_i - a_i| \leq \delta$, then $\rho(F, G) \leq \delta$.

We now verify the statement surrounding (C.19). Since the difference of two equicontinuous functions is equicontinuous, it suffices to verify the statement with $\hat{D}_{n,1}(\mathbf{X})$ replaced by $\hat{P}_{n,1}(\mathbf{X})$. This entails showing that, given $\delta > 0$, there exists $\delta' > 0$ such that

$$\sqrt{n}|\frac{\sum_{i=1}^{n-1} X_{\pi(i)}X_{\pi(i+1)}}{\sum_{i=1}^{n-1} X_{\pi(i)}} - \frac{\sum_{i=1}^{n-1} X'_{\pi(i)}X'_{\pi(i+1)}}{\sum_{i=1}^{n-1} X'_{\pi(i)}}| \tag{C.20}$$

is $\leq \delta$ if $|h_n - h| \leq \delta'$. Let $T = \sum_{i=1}^{n-1} X_{\pi(i)}$ and $T' = \sum_{i=1}^{n-1} X'_{\pi(i)}$. Clearly,

$$S(h_n) - 1 \leq T \leq S(h_n)$$

and

$$S(h) - 1 \leq T' \leq S(h).$$

Now, (C.20) can be expressed as

$$\frac{\sqrt{n}}{T'T}(T' - T)\sum_{i=1}^{n-1} X_{\pi(i)}X_{\pi(i+1)} + \frac{\sqrt{n}}{T'}\sum_{i=1}^{n-1}\left[X_{\pi(i)}X_{\pi(i+1)} - X'_{\pi(i)}X'_{\pi(i+1)}\right]. \tag{C.21}$$

Separate (C.21) into the terms $A_n + B_n$, and we show each term is $\leq \delta/2$ for an appropriate choice of δ'. The sum in A_n can be bounded by T, so that

$$|A_n| \leq \sqrt{n}|\frac{T - T'}{T'}|.$$

Consider the case where $h_n > h$, so then

$$\begin{aligned}|A_n| &\leq \sqrt{n}\left(\frac{\lfloor \frac{n}{2} + \sqrt{n}h_n \rfloor - \lfloor \frac{n}{2} + \sqrt{n}h \rfloor - 1}{\lfloor \frac{n}{2} + \sqrt{n}h \rfloor - 1}\right) \\ &\leq \sqrt{n}\left(\frac{\frac{n}{2} + \sqrt{n}h_n - \left(\frac{n}{2} + \sqrt{n}h - 1\right) - 1}{\frac{n}{2} + \sqrt{n}h - 2}\right) \\ &= \frac{h_n - h}{\frac{1}{2} + \frac{h}{\sqrt{n}} - \frac{2}{\sqrt{n}}}.\end{aligned}$$

In general,

$$|A_n| \leq \frac{|h_n - h|}{\frac{1}{2} + \frac{h}{\sqrt{n}} - \frac{2}{\sqrt{n}}},$$

which can be made to be $\leq \delta$ for all large n for δ' chosen sufficiently small, as long as h is restricted to be in a bounded set.

To bound B_n, note that by the coupling construction of $\mathbf{X}$ and $\mathbf{X}'$, $\mathbf{X}$ and $\mathbf{X}'$ differ in at most $|S(h_n) - S(h)|$ entries. Therefore, for any π, $X_{\pi(i)}X_{\pi(i+1)} - X'_{\pi(i)}X'_{\pi(i+1)}$ are 0except for at most $2|S(h_n) - S(h)|$ number of them, and the nonzero ones can be bounded above by 1. Therefore,

$$|B_n| \leq \frac{2\sqrt{n}}{T'}|S(h_n) - S(h)|,$$

which, similar to A_n, can be bounded above by

$$|B_n| \leq \frac{2|h_n - h| + \frac{2}{\sqrt{n}}}{\frac{1}{2} + \frac{h}{\sqrt{n}}}.$$

Therefore, for large enough n, chosen independently of π, the bound can be made $\leq \delta/2$ if $|h_n - h| \leq \delta'$ for sufficiently small δ'.

In summary, for some sufficiently chosen positive δ' and large enough n, (C.19) holds for all π, and the lemma follows.

We now turn to the proof of the main result. Assume the opposite. Then, there exists $\epsilon > 0$ such that

$$\rho(L_n(h), N(0,1)) > \epsilon$$

for infinitely many n. Assume this holds for all large n, or apply the argument below to a subsequence. Let $\delta = \epsilon/2$. Then, there exists $\delta' > 0$ such that for large n, $\rho(L_n(h), L_n(h_n)) \leq \delta$ if $|h_n - h| \leq \delta'$. Let E_n be the set of $S_n(h_n)$ with $|h_n - h| \leq \delta'$. Consider i.i.d. sampling with $p = \frac{1}{2}$ and let $\hat{S}_n$ be the number of successes. Then,

$$\mathbb{P}_{p=\frac{1}{2}}\left\{\hat{S}_n \in E_n\right\} = \mathbb{P}_{p=\frac{1}{2}}\left\{n^{-1/2}|S_n - \frac{n}{2}| \leq \delta'\right\} \to c > 0. \tag{C.22}$$

Let $\hat{h}_n = n^{1/2}\left(\hat{S}_n - \frac{n}{2}\right)$. When E_n occurs, we have for sufficiently large n,

$$\rho\left(L_n\left(\hat{h}_n\right), L_n\left(h\right)\right) \leq \delta = \epsilon/2,$$

which by the triangle inequality implies

$$\rho\left(L_n\left(\hat{h}_n\right), L_n\left(h\right)\right) \geq \epsilon/2$$

for sufficiently large n. Note $L_n\left(\hat{h}_n\right)$ is indeed the (random) permutation distribution based on i.i.d. Bernoulli trials with success probability $1/2$. But, because convergence in the Levy metric is weaker than convergence of distributions in the supremum metric,

$$\mathbb{P}_{p=\frac{1}{2}}\left\{\rho\left(L_n\left(\hat{h}_n\right), N\left(0,1\right)\right) \geq \epsilon/2\right\} \to 0,$$

which is a contradiction because of the probability of E_n does not tend to 0, by (C.22).

C.11 Proof of Theorem 3.6

To show (3.9), it suffices to show, given any subsequence n' there exists a further subsequence n'' such that $\sup_t |\hat{R}_{n''}\left(t\right) - \Phi\left(t\right)| \to 0$ with probability one. Now appeal to the Almost Sure Representation Theorem, and construct $\tilde{S}_n$ with the same distribution as $\hat{S}_n$ such that $n^{-1/2}\left(\tilde{S}_n - p\right)$ converges to some H almost surely. But, for every sequence for which $n^{-1/2}\left(\tilde{S}_n - p\right)$ converges, we have $\sup_t |\hat{R}_{n''}\left(t\right) - \Phi\left(t\right)| \to 0$, by Theorem 3.5. Since convergence occurs with probability one, the result holds.

D Second Order Approximations

Theorem D.1. *Under the assumption that* $\mathbf{X} = \{X_i\}_{i=1}^n$ *is a sequence of independent and identically distributed Bernoulli(p) random variables, then*
(i) the expectation of $\hat{P}_k(\mathbf{X})$ *has a second order approximation given by*

$$\mathbb{E}\left[\hat{P}_k(\mathbf{X})\right] = p + n^{-1}p\left(1 - p^{-k}\right) + O\left(n^{-2}\right). \tag{D.1}$$

(ii) the expectation of $\hat{D}_k(\mathbf{X})$ has a second order approximation given by

$$\mathbb{E}\left[\hat{D}_k(\mathbf{X})\right] = n^{-1}\left(1 - (1-p)^{1-k} - p^{1-k}\right) + O\left(n^{-2}\right). \tag{D.2}$$

(iii) $\mathrm{Cov}\left(\hat{P}_k(\mathbf{X}), \hat{Q}_k(\mathbf{X})\right)$ *has a second order approximation given by*

$$\mathrm{Cov}\left(\hat{P}_k(\mathbf{X}), \hat{Q}_k(\mathbf{X})\right) = O\left(n^{-2}\right). \tag{D.3}$$

Figure 9 displays the second order approximation to the expectations of $\hat{P}_k(\mathbf{X})$ and $\hat{D}_k(\mathbf{X})$ under H_0 for $k = 1$ and 3 and $p \in (0.25, 0.5, 0.75)$. The approximation is not suitable for small values of p and n for large k, but does a remarkably good job for moderate values of p and n. Note that the second order approximation for $\hat{D}_k(\mathbf{X})$ when $k = 1$ is the same for all p, and in fact, the Monte Carlo estimates for the finite sample values of $\hat{D}_k(\mathbf{X})$ when $k = 1$ are too close to discern at the scale that we have plotted them.

D.1 Proof of Theorem D.1 (i)

For notational simplicity, let $\bar{V}_k = n^{-1}\sum_{i=1}^{n-k} Y_{i,k}$. Let $g(\theta_1, \theta_2) = \theta_1/\theta_2$. The Taylor expansion of $\mathbb{E}\left[\hat{P}_{n,k}(\mathbf{X})\right] = \mathbb{E}\left[g\left(\bar{V}_k, \bar{V}_{k-1},\right)\right]$ about $\left(p^{k+1}, p^k\right)$ is given by

$$\begin{aligned}
\mathbb{E}\left[\hat{P}_{n,k}(\mathbf{X})\right] &= g\left(p^{k+1}, p^k\right) + \frac{1}{2}\operatorname{Var}\left(\bar{V}_k\right)\frac{\partial^2 g}{\partial \bar{V}_k^2}\left(p^{k+1}, p^k\right) \\
&\quad + \frac{1}{2}\operatorname{Var}\left(\bar{V}_{k-1}\right)\frac{\partial^2 g}{\partial \bar{V}_{k-1}^2}\left(p^{k+1}, p^k\right) + \operatorname{Cov}\left(\bar{V}_k, \bar{V}_{k-1}\right)\frac{\partial^2 g}{\partial \bar{V}_k \partial \bar{V}_{k-1}}\left(p^{k+1}, p^k\right) + O\left(n^{-2}\right). \\
&= p + p^{1-2k}\operatorname{Var}\left(\bar{V}_{k-1}\right) - \frac{1}{p^{2k}}\operatorname{Cov}\left(\bar{V}_k, \bar{V}_{k-1}\right) + O\left(n^{-2}\right).
\end{aligned}$$

This is given by

$$\begin{aligned}
\mathbb{E}\left[\hat{P}_{n,k}(\mathbf{X})\right] &= p + \frac{p^{1-2k}}{n}\left(p^k - (2k-1)\,p^{2k} + \frac{2p^{k+1} - 2p^{2k}}{1-p}\right) + \\
&\quad - \frac{1}{np^{2k}}\left(2p^{k+1} - 2kp^{2k+1} + \frac{2p^{k+2} - 2p^{2k+1}}{1-p}\right) + O\left(n^{-2}\right) \\
&= p + n^{-1}p\left(1 - p^{-k}\right) + O\left(n^{-2}\right).
\end{aligned}$$

D.2 Proof of Theorem D.1 (ii)

From Proposition 3, we can see that

$$\begin{aligned}
\mathbb{E}\left[\hat{D}_{n,k}(\mathbf{X})\right] &= \mathbb{E}\left[\hat{P}_{n,k}(\mathbf{X})\right] - \left(1 - \mathbb{E}\left[\hat{Q}_{n,k}(\mathbf{X})\right]\right) \\
&= p + n^{-1}p\left(1 - p^{-k}\right) - \left(1 - (1-p) - n^{-1}(1-p)\left(1 - (1-p)^{-k}\right)\right) + O\left(n^{-2}\right) \\
&= n^{-1}\left(1 - (1-p)^{1-k} - p^{1-k}\right) + O\left(n^{-2}\right).
\end{aligned}$$

D.3 Proof of Theorem D.1 (iii)

For notational simplicity, let $\bar{V}_k = n^{-1}\sum_{i=1}^{n-k} Y_{i,k}$ and $\bar{W}_k = n^{-1}\sum_{i=1}^{n-k} Z_{i,k}$. Let $g(\theta_1, \theta_2, \theta_3, \theta_4) = \theta_1\theta_3/\theta_2\theta_4$. The Taylor expansion of $\mathbb{E}\left[\hat{P}_{n,k}(\mathbf{X})\hat{Q}_{n,k}(\mathbf{X})\right] = \mathbb{E}\left[g\left(\bar{V}_k, \bar{V}_{k-1}, \bar{W}_k, \bar{W}_{1-k}\right)\right]$ about $\rho = \left(p^{k+1}, p^k, (1-p)^{k+1}, (1-p)^k\right)$ is given by

$$
\begin{aligned}
&\mathbb{E}\left[\hat{P}_{n,k}(\mathbf{X})\hat{Q}_{n,k}(\mathbf{X})\right] \\
&= g\left(\rho\right) \\
&+ \frac{1}{2}\operatorname{Var}\left(\bar{V}_k\right)\frac{\partial^2 g}{\partial \bar{W}_k^2}\left(\rho\right) + \frac{1}{2}\operatorname{Var}\left(\bar{V}_{k-1}\right)\frac{\partial^2 g}{\partial \bar{W}_{k-1}^2}\left(\rho\right) \\
&+ \frac{1}{2}\operatorname{Var}\left(\bar{W}_k\right)\frac{\partial^2 g}{\partial \bar{V}_k^2}\left(\rho\right) + \frac{1}{2}\operatorname{Var}\left(\bar{W}_{k-1}\right)\frac{\partial^2 g}{\partial \bar{V}_{k-1}^2}\left(\rho\right) \\
&+ \operatorname{Cov}\left(\bar{V}_k, \bar{V}_{k-1}\right)\frac{\partial^2 g}{\partial \bar{V}_k \partial \bar{V}_{k-1}}\left(\rho\right) + \operatorname{Cov}\left(\bar{V}_k, \bar{W}_k\right)\frac{\partial^2 g}{\partial \bar{V}_k \partial \bar{W}_k}\left(\rho\right) \\
&+ \operatorname{Cov}\left(\bar{V}_k, \bar{W}_{k-1}\right)\frac{\partial^2 g}{\partial \bar{V}_k \partial \bar{W}_{k-1}}\left(\rho\right) + \operatorname{Cov}\left(\bar{V}_{k-1}, \bar{W}_k\right)\frac{\partial^2 g}{\partial \bar{V}_{k-1} \partial \bar{W}_k}\left(\rho\right) \\
&\left(\rho\right) + \operatorname{Cov}\left(\bar{V}_{k-1}, \bar{W}_{k-1}\right)\frac{\partial^2 g}{\partial \bar{V}_{k-1} \partial \bar{W}_{k-1}}\left(\rho\right) + \operatorname{Cov}\left(\bar{W}_k, \bar{W}_{k-1}\right)\frac{\partial^2 g}{\partial \bar{W}_k \partial \bar{W}_{k-1}}\left(\rho\right) + O\left(n^{-2}\right) \\
&= p\left(1-p\right) + p^{1-2k}\left(1-p\right)\operatorname{Var}\left(\bar{V}_{k-1}\right) + p\left(1-p\right)^{1-2k}\operatorname{Var}\left(\bar{W}_{k-1}\right) \\
&- \frac{\left(1-p\right)}{p^{2k}}\operatorname{Cov}\left(\bar{V}_k, \bar{V}_{k-1}\right) + \frac{1}{p^k\left(1-p\right)^k}\operatorname{Cov}\left(\bar{V}_k, \bar{W}_k\right) - \frac{\left(1-p\right)^{1-k}}{p^k}\operatorname{Cov}\left(\bar{V}_k, \bar{W}_{k-1}\right) \\
&- \frac{p^{1-k}}{\left(1-p\right)^k}\operatorname{Cov}\left(\bar{V}_{k-1}, \bar{W}_k\right) + p^{1-k}\left(1-p\right)^{1-k}\operatorname{Cov}\left(\bar{V}_{k-1}, \bar{W}_{k-1}\right) \\
&- \frac{p}{\left(1-p\right)^{2k}}\operatorname{Cov}\left(\bar{W}_k, \bar{W}_{k-1}\right) + O\left(n^{-2}\right)
\end{aligned}
$$

This is given by

$$
\begin{aligned}
&\mathbb{E}\left[\hat{P}_{n,k}(\mathbf{X})\hat{Q}_{n,k}(\mathbf{X})\right] \\
&= p\left(1-p\right) + \frac{p^{1-2k}\left(1-p\right)}{n}\left(p^k - \left(2k-1\right)p^{2k} + \frac{2p^{k+1} - 2p^{2k}}{1-p}\right) \\
&+ \frac{p\left(1-p\right)^{1-2k}}{n}\left(\left(1-p\right)^k - \left(2k-1\right)\left(1-p\right)^{2k} + \frac{2(1-p)^{k+1} - 2(1-p)^{2k}}{p}\right) \\
&- \frac{\left(1-p\right)}{np^{2k}}\left(2p^{k+1} - 2kp^{2k+1} + \frac{2p^{k+2} - 2p^{2k+1}}{1-p}\right) \\
&+ \frac{1}{np^k\left(1-p\right)^k}\left(-\left(2k+1\right)p^{k+1}\left(1-p\right)^{k+1}\right) - \frac{\left(1-p\right)^{1-k}}{np^k}\left(-\left(2k\right)p^{k+1}\left(1-p\right)^k\right) \\
&- \frac{p^{1-k}}{n\left(1-p\right)^k}\left(-\left(2k\right)p^k\left(1-p\right)^{k+1}\right) + \frac{p^{1-k}\left(1-p\right)^{1-k}}{n}\left(-\left(2k-1\right)p^k\left(1-p\right)^k\right) \\
&- \frac{p}{n\left(1-p\right)^{2k}}\left(2(1-p)^{k+1} - 2k(1-p)^{2k+1} + \frac{2(1-p)^{k+2} - 2(1-p)^{2k+1}}{p}\right) + O\left(n^{-2}\right) \\
&= p\left(1-p\right) - n^{-1}\left(1-p\right)^{1-k}p^{1-k}\left(p^k + \left(1-p\right)^k\left(1-2p^k\right)\right) + O\left(n^{-2}\right).
\end{aligned}
$$

Therefore, from Theorem 4, we can see that

$$\begin{aligned}
&\operatorname{Cov}\left(\hat{P}_{n,k}(\mathbf{X}), \hat{Q}_{n,k}(\mathbf{X})\right) \\
&= \mathbb{E}\left[\hat{P}_{n,k}(\mathbf{X})\hat{Q}_{n,k}(\mathbf{X})\right] - \mathbb{E}\left[\hat{P}_{n,k}(\mathbf{X})\right]\mathbb{E}\left[\hat{Q}_{n,k}(\mathbf{X})\right] \\
&= p\,(1-p) - n^{-1}\left((1-p)^{1-k}\,p^{1-k}\left(p^{k} + (1-p)^{k}\left(1-2p^{k}\right)\right)\right) + O\left(n^{-2}\right) \\
&- p\,(1-p) - n^{-1}\left(p\,(1-p)\left(2 - p^{-k} - (1-p)^{-k}\right)\right) \\
&+ n^{-2}\left(p\,(1-p)\left(1-p^{-k}\right)\left(1-(1-p)^{-k}\right)\right) \\
&= n^{-2}\left(p\,(1-p)\left(1-p^{-k}\right)\left(1-(1-p)^{-k}\right)\right) + O\left(n^{-2}\right).
\end{aligned}$$

E Normal Approximation Testing Procedures

Theorem E.1. *Under the assumption that $\mathbf{X} = \{X_i\}_{i=1}^{n}$ is a sequence of independent and identically distributed Bernoulli(p) random variables, then the ratio*

$$\frac{s_{p,i}^{2}/V_{i,k}}{n^{-1}\hat{P}_{n,k}\left(\mathbf{X}_i\right)^{1-k}\left(1-\hat{P}_{n,k}\left(\mathbf{X}_i\right)\right)}$$

tends to 1 *in probability.*

Theorem E.2. *Let $\mathbf{X} = \{X_i\}_{i=1}^{n}$ be a sequence of independent and identically distributed Bernoulli(p_i) random variables, then the ratio of*

$$\left(\frac{\left(V_{ik}-1\right)s_{p,i}^{2} + \left(W_{ik}-1\right)s_{q,i}^{2}}{V_{ik}+W_{ik}-2}\right)\left(\frac{1}{V_{ik}} + \frac{1}{W_{ik}}\right) \tag{E.1}$$

and the asymptotic variance of $\hat{D}_{n,k}(\mathbf{X}_i)$, given by (6)*, tends to* 1 *in probability.*

E.1 Proof of Theorem E.1

Let $V_k = \sum_{i=1}^{n-k} Y_{i,k}$. Miller and Sanjurjo (2018) estimate of the variance of $\hat{P}_k(\mathbf{X})$ as

$$s_{p,i}^{2} = \frac{\hat{P}_{n,k}\left(\mathbf{X}\right)\left(1-\hat{P}_{n,k}\left(\mathbf{X}\right)\right)}{V_k}.$$

Since by Theorem 2.1, $\hat{P}_{n,k}(\mathbf{X})$ is a consistent estimator of p, it follows that

$$\frac{\hat{P}_{n,k}\left(\mathbf{X}\right)^{1-k}\left(1-\hat{P}_{n,k}\left(\mathbf{X}\right)\right)/n}{p^{1-k}\left(1-p\right)/n} \xrightarrow{P} 1.$$

So, in order to show

$$\frac{\hat{P}_{n,k}\left(\mathbf{X}\right)\left(1-\hat{P}_{n,k}\left(\mathbf{X}\right)\right)/V_k}{\hat{P}_{n,k}\left(\mathbf{X}\right)^{1-k}\left(1-\hat{P}_{n,k}\left(\mathbf{X}\right)\right)/n} \xrightarrow{P} 1$$

it will suffice to to show that

$$\frac{\hat{P}_{n,k}\left(\mathbf{X}\right)\left(1-\hat{P}_{n,k}\left(\mathbf{X}\right)\right)/V_k}{p\left(1-p\right)/np^k} \xrightarrow{P} 1.$$

This is equivalent to

$$\frac{np^k}{V_k} \xrightarrow{P} 1$$

and in turn to

$$\frac{V_k}{np^k} \xrightarrow{P} 1.$$

This follow from

$$\frac{\mathbb{E}\left[V_k\right]}{np^k} = \frac{\left(n-k\right)p^k}{np^k} \to 1$$

and

$$\operatorname{Var}\left(\frac{V_k}{np^k}\right) = \frac{1}{n^2p^{2k}}\operatorname{Var}\left(V_k\right) = O\left(n^{-1}\right) \to 0$$

as $\operatorname{Var}\left(W_k\right)$ is given by

$$n\left(p^k - \left(2k-1\right)p^{2k} + \frac{2p^{k+1}-2p^{2k}}{1-p}\right).$$

E.2 Proof of Theorem E.2

By the proof of Theorem E.1, the ratio of

$$\left(\frac{\left(V_{ik}-1\right)s_{p,i}^2 + \left(W_{ik}-1\right)s_{q,i}^2}{V_{ik}+W_{ik}-2}\right)\left(\frac{1}{V_{ik}}+\frac{1}{W_{ik}}\right)$$

and

$$
\begin{aligned}
&\left(\frac{np^{k+1}\left(1-p\right)+n\left(1-p\right)^{k+1}p}{np^{k}+n\left(1-p\right)^{k}}\right)\left(\frac{1}{np^{k}}+\frac{1}{n\left(1-p\right)^{k}}\right)\\
&=\frac{p\left(1-p\right)\left(\left(1-p\right)^{k}+p^{k}\right)}{np^{k}\left(1-p\right)^{k}}\\
&=n^{-1}\left(p\left(1-p\right)\right)^{1-k}\left(\left(1-p\right)^{k}+p^{k}\right)
\end{aligned}
$$

tends to 1 in probability as n grows to infinity.

References

Bradley, R.C., 2005. Basic properties of strong mixing conditions. A survey and some open questions. *Probability Surveys*, 2, pp.107-144.

Chung, E. and Romano, J.P., 2013. Exact and asymptotically robust permutation tests. *The Annals of Statistics*, 41(2), pp.484-507.

Ibragimov, I.A., 1962. Some limit theorems for stationary processes. *Theory of Probability & Its Applications*, 7(4), pp.349-382.

Miller, J.B. and Sanjurjo, A., 2018. Surprised by the hot hand fallacy? A truth in the law of small numbers. Econometrica, 86(6), pp.2019-2047.

Politis, D.N., Romano, J.P. and Wolf, M., 1999. *Subsampling*. Springer-Verlag, NY.

Rinott, Y., 1994. On normal approximation rates for certain sums of dependent random variables. *Journal of Computational and Applied Mathematics*, 55(2), pp.135-143.

Stein, C., 1986. *Approximate Computation of Expectations*. IMS. Hayward, CA.

Wald, A. and Wolfowitz, J., 1940. On a test whether two samples are from the same population. *The Annals of Mathematical Statistics*, 11(2), pp.147-162.