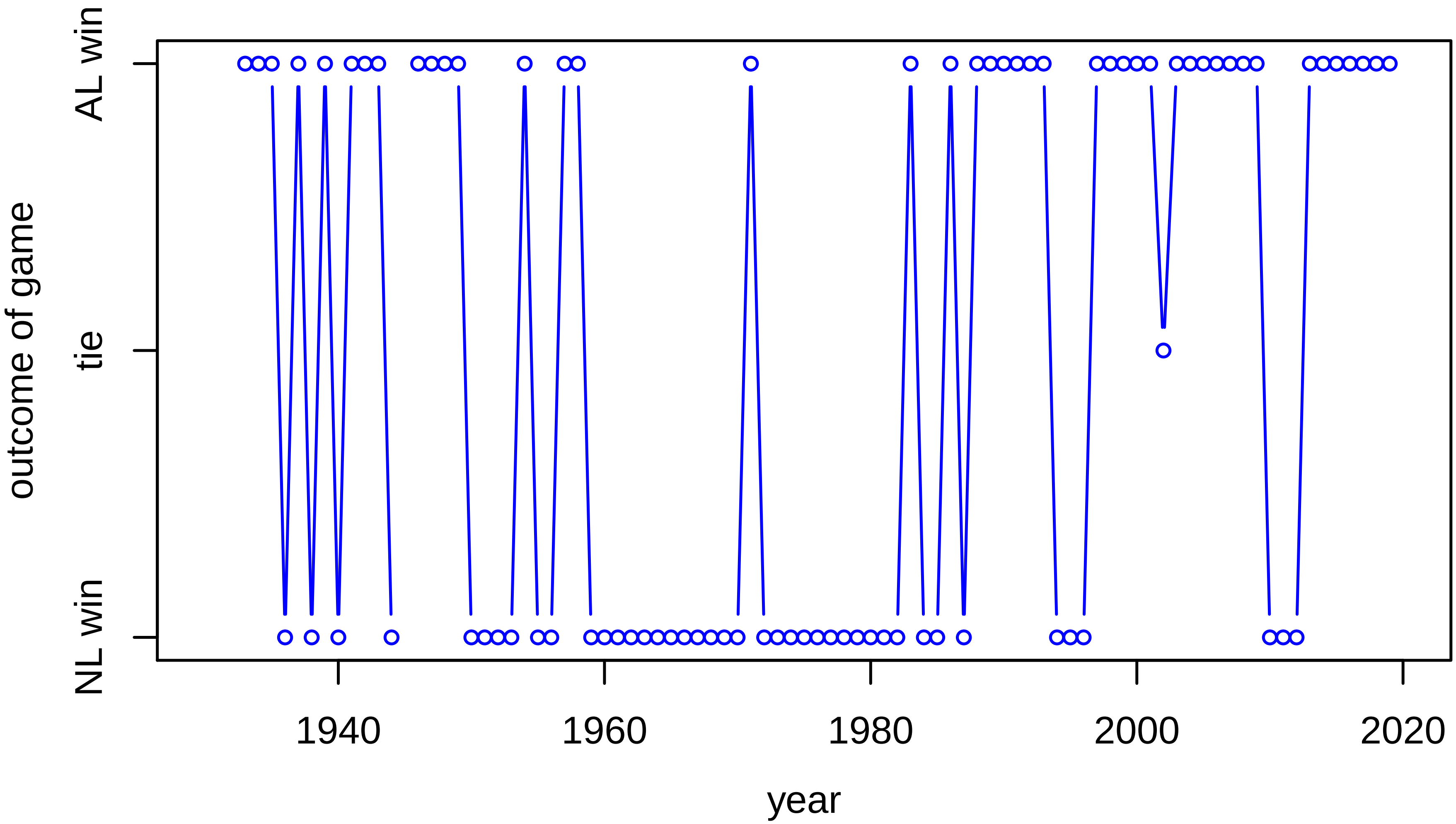
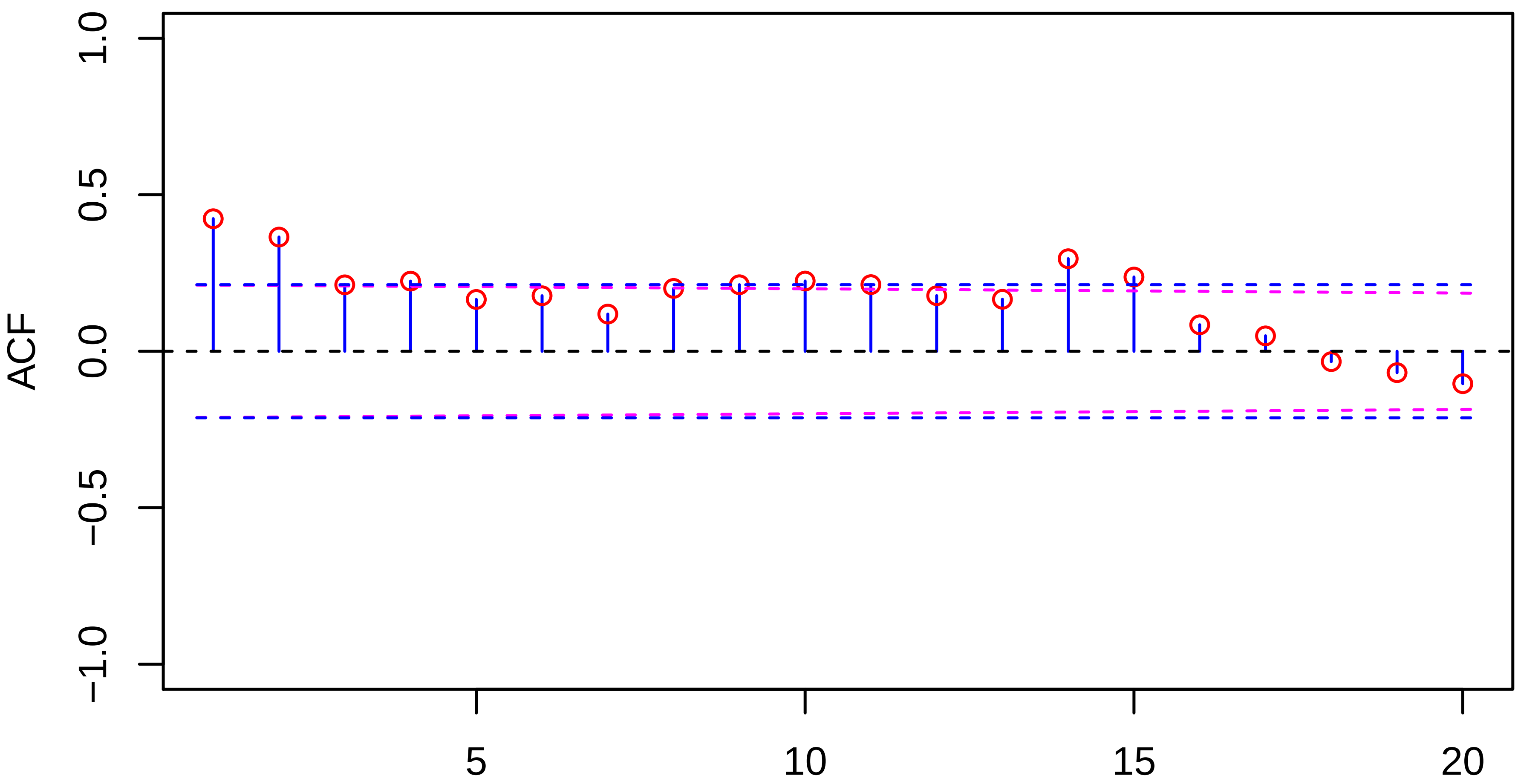
Testing for IID Noise/White Noise: I

- want to be able to test null hypothesis that time series $\{x_t\}$ or set of residuals $\{r_t\}$ is $\text{IID}(0, \sigma^2)$ or $\text{WN}(0, \sigma^2)$
- many such tests exist, including
  - informal test based on sample ACF (see overhead II–66)
  - portmanteau test
  - turning point test
  - difference-sign test
  - rank test
  - runs test
- will illustrate tests using all-star game time series $\{x_t\}$ & residuals $\{z_t\}$ from AR(1) model for detrended Lake Huron levels

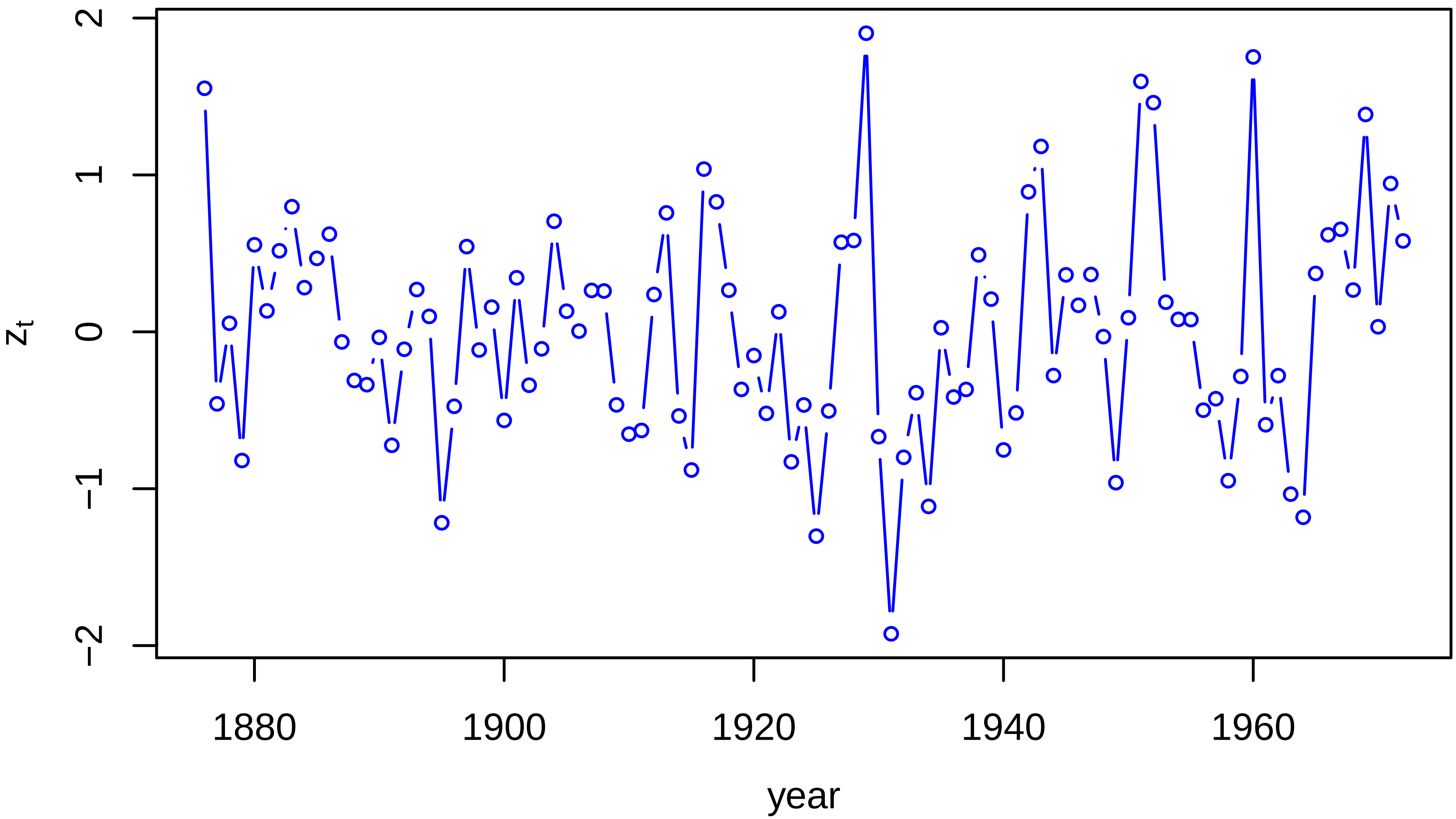
First All-Star Baseball Game in Each Year (I–22)



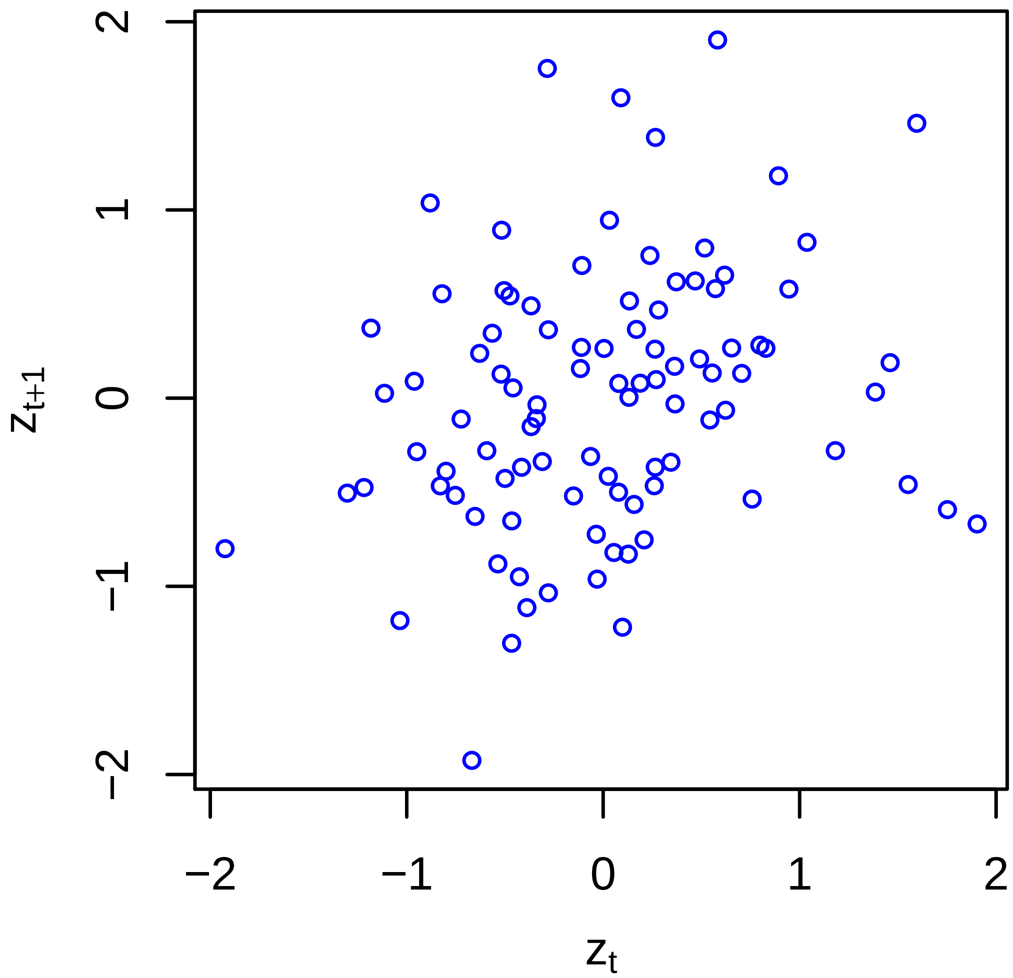
Sample ACF for All-Star Game $\{x_t\}$



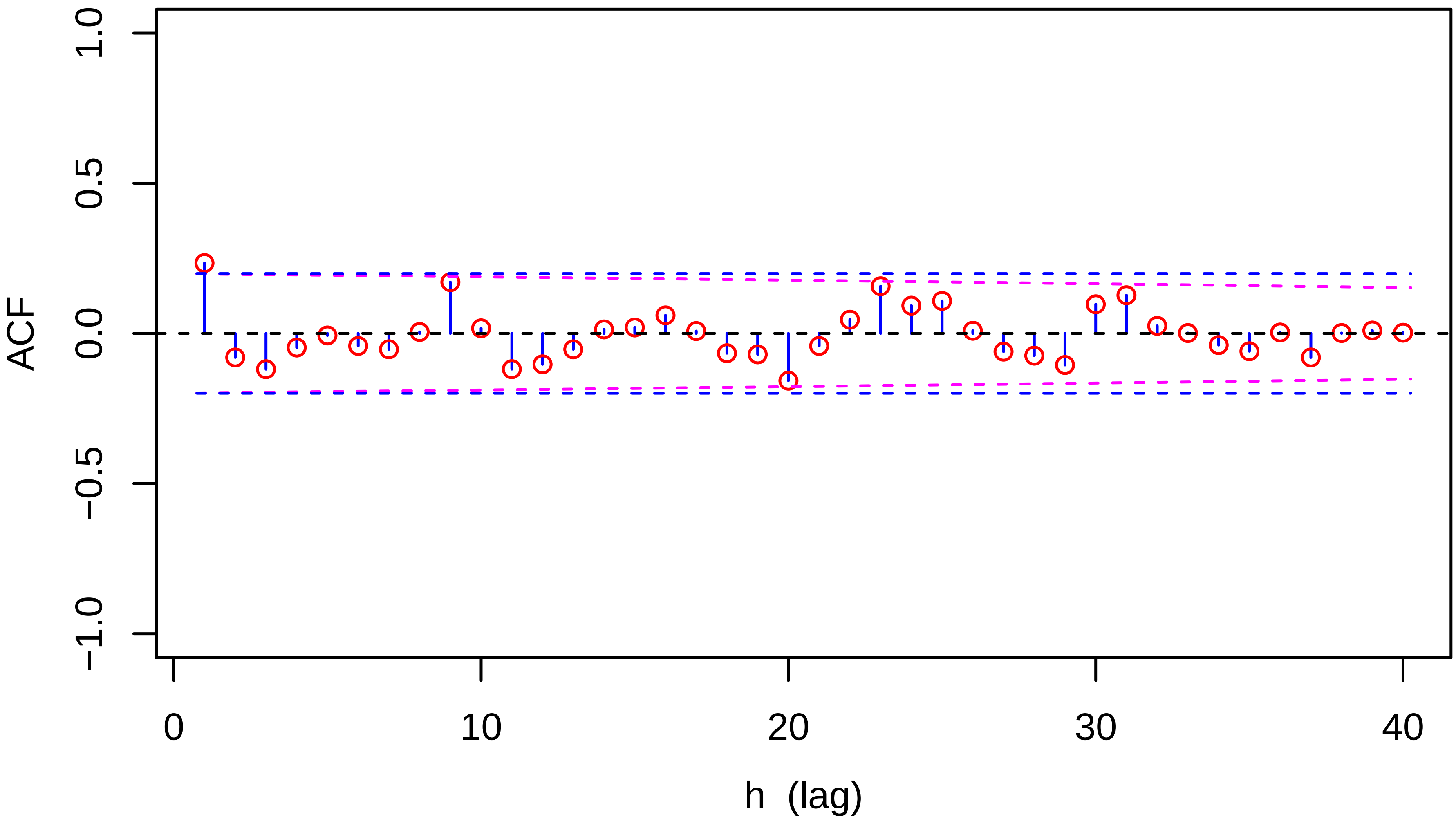
AR(1) Residuals $z_t = r_t - \hat{\phi} r_{t-1}$ (III–15)



Unit Lag Scatter Plot of AR(1) Residuals z_t (III–16)



Sample ACF for AR(1) Residuals z_t (III–17)



Prelude to Portmanteau Test

- Q: what exactly does ‘portmanteau’ mean?
- as a noun: either

  large trunk or suitcase, typically made of stiff leather and opening into two equal parts

  or

  word blending sounds and combining meanings of two others, for example motel (from motor and hotel)

- as an adjective:

  consisting of or combining two or more separable aspects or qualities: a portmanteau movie composed of excerpts from his/her/their most famous films

  (definitions from Dictionary app on Mac OS X Yosemite)

Portmanteau Test: I

- can regard as formal version of informal test based on sample ACF, which makes use of fact that, when $\{X_t\} \sim \mathrm{IID}(\mu, \sigma^2)$, then corresponding $\hat{\rho}_X(1), \ldots, \hat{\rho}_X(h)$ are approximately IID $\mathcal{N}(0, 1/n)$ for large sample sizes n (but: h/n must be 'small'!)
- hence
$$\frac{\hat{\rho}_X(j)}{\sqrt{1/n}} = n^{1/2}\hat{\rho}_X(j), \quad j = 1, \ldots, h,$$
are approximately IID standard normal RVs (i.e., $\mathcal{N}(0, 1)$)
- recall that, if $Z_1, Z_2, \ldots, Z_\nu$ are IID $\mathcal{N}(0, 1)$, then
$$Z_1^2 + Z_2^2 + \cdots + Z_\nu^2 \stackrel{\text{def}}{=} \chi^2_\nu$$
obeys a chi-square distribution with ν degrees of freedom

Portmanteau Test: II

- hence, under null hypothesis $\{X_t\} \sim \text{IID}(\mu, \sigma^2)$,

$$Q(h) \stackrel{\text{def}}{=} \sum_{j=1}^{h} \left(n^{1/2}\hat{\rho}_X(j)\right)^2 = n\sum_{j=1}^{h}\hat{\rho}_X^2(j)$$

  should obey a χ_h^2 distribution
- $Q(h)$ will be large under alternative hypotheses, so can reject null hypothesis at level of significance α if $Q(h) > \chi_{1-\alpha}^2(h)$, where $\chi_{1-\alpha}^2(h)$ is $(1-\alpha)$th quantile of χ_h^2 distribution
- refinement is Ljung–Box version of portmanteau test:

$$Q_{\text{LB}}(h) = n(n+2)\sum_{j=1}^{h}\frac{\hat{\rho}_X^2(j)}{n-j}$$

  (obeys χ_h^2 distribution to better approximation than $Q(h)$ does)

Portmanteau Test: III

- if we test residuals $\{r_t\}$ rather than a time series $\{x_t\}$, then

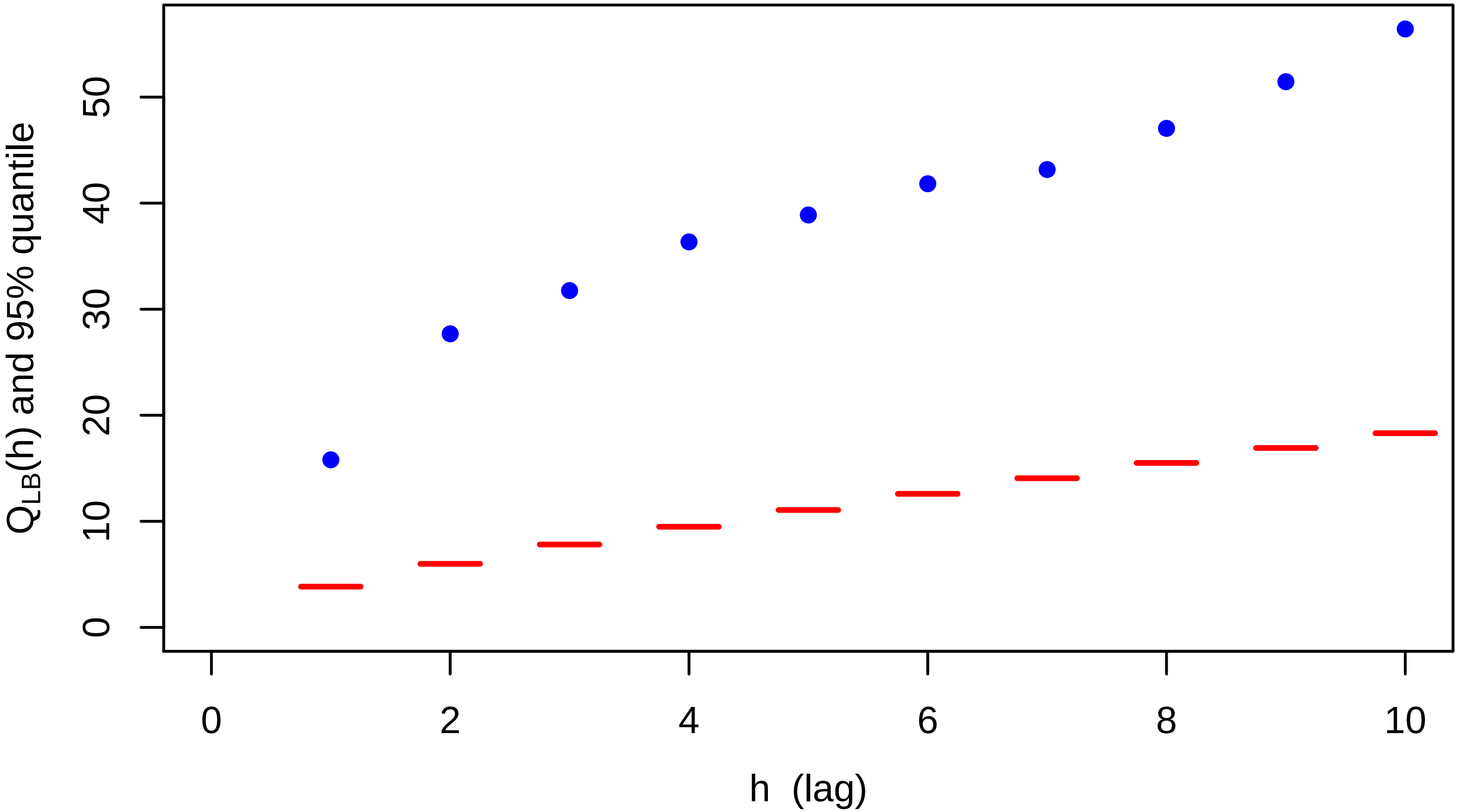
$$Q_{\text{LB}}(h) = n(n+2)\sum_{j=1}^{h}\frac{\hat{\rho}_R^2(j)}{n-j}$$

  has the same form as before, but now obeys a χ^2_{h-p} distribution, where p is the number of parameters estimated in forming $\{r_t\}$

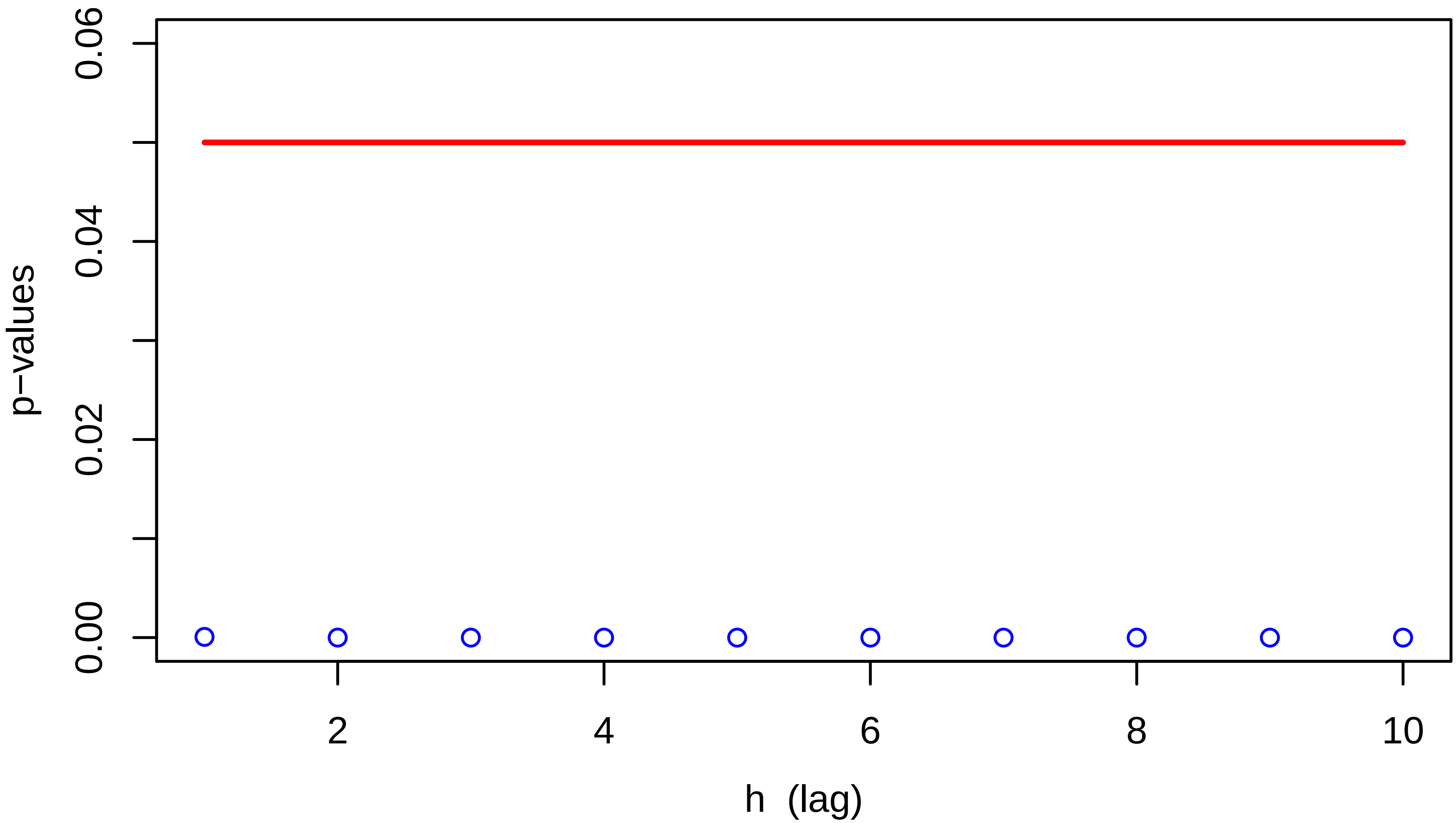
- as examples, let's look at $Q_{\text{LB}}(h)$ for all-star game time series $\{x_t\}$ and Lake Huron residuals $\{z_t\}$, for which $p = 3$ arguably (have estimated intercept, slope and the AR(1) parameter ϕ)
  - will focus on tests with level $\alpha = 0.05$
  - red horizontal lines show $\chi^2_{0.95}(h-p)$ (note: $p = 0$ for all-star game series)
  - blue dots show $Q_{\text{LB}}(h)$

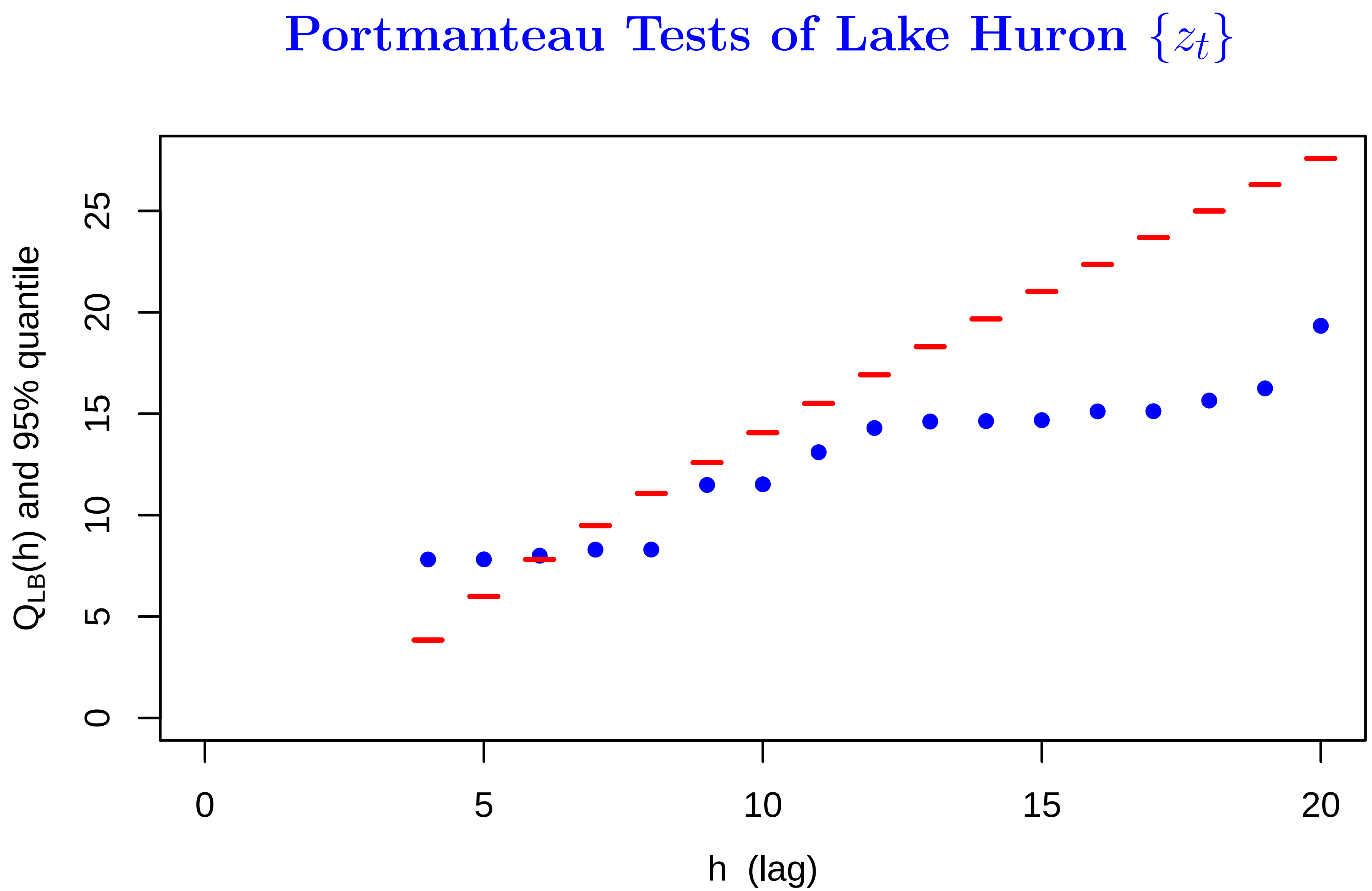
CC: 183, SS: 139

Portmanteau Tests of All-Star Game $\{x_t\}$

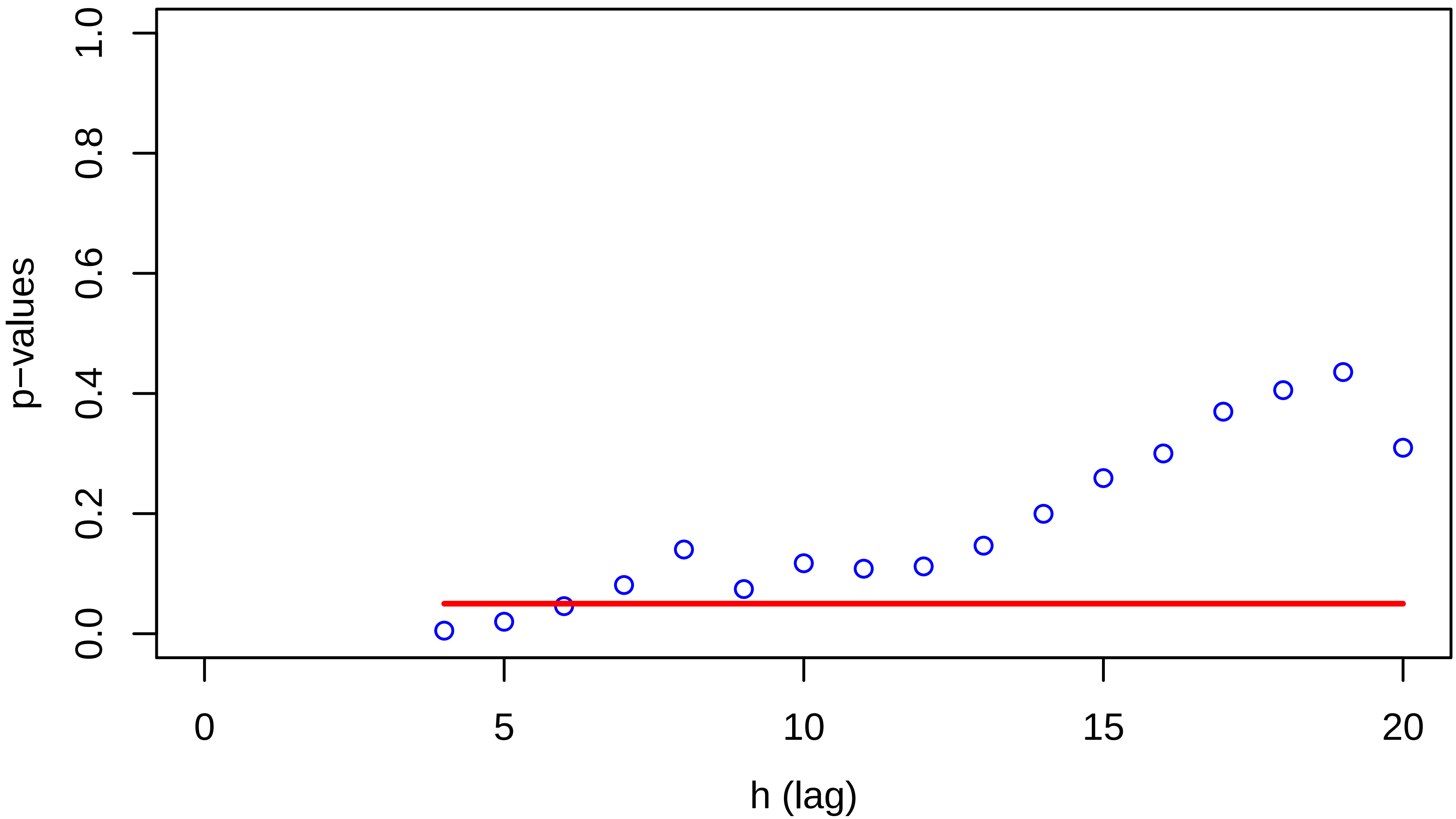


p-values for Portmanteau Tests of All-Star Game $\{x_t\}$




Portmanteau Tests of Lake Huron $\{z_t\}$
$Q_{LB}(h)$ and 95% quantile
0
5
10
15
20
25
0
5
10
15
20
h (lag)

p-values for Portmanteau Tests of Lake Huron $\{z_t\}$



Turning Point Test: I

- time series $\{x_t\}$ is said to have a turning point at t if one of the following two patterns exist:
  1. $x_t > x_{t-1}$ and $x_t > x_{t+1}$
  2. $x_t < x_{t-1}$ and $x_t < x_{t+1}$
- if X_{t-1}, X_t and X_{t+1} are IID RVs with a *continuous* distribution, the following six events are equally likely:

$$\begin{array}{ll}
X_{t-1} < X_t < X_{t+1} & \text{(no turning point)} \\
X_{t-1} < X_{t+1} < X_t & \text{(turning point)} \\
X_t < X_{t-1} < X_{t+1} & \text{(turning point)} \\
X_t < X_{t+1} < X_{t-1} & \text{(turning point)} \\
X_{t+1} < X_{t-1} < X_t & \text{(turning point)} \\
X_{t+1} < X_t < X_{t-1} & \text{(no turning point)}
\end{array}$$

Turning Point Test: II

- let T be number of turning points in IID sequence of length n
- since probability of a turning point is $2/3$ under IID hypothesis,
$$\mu_T \stackrel{\text{def}}{=} E\{T\} = 2(n-2)/3$$
- can show that, under IID hypothesis,
$$\sigma_T^2 \stackrel{\text{def}}{=} \operatorname{var}\{T\} = \frac{16n-29}{90}$$
- large $|T - \mu_T|$ indicates time series is either fluctuating more rapidly than what would be expected from an IID sequence or there are fewer fluctuations than expected
- under IID hypothesis, T is approximately $\mathcal{N}(\mu_T, \sigma_T^2)$, so can reject null at level α if $|T - \mu_T|/\sigma_T > \Phi_{1-\alpha/2}$, where $\Phi_{1-\alpha/2}$ is $1 - \alpha/2$ quantile for standard normal distribution

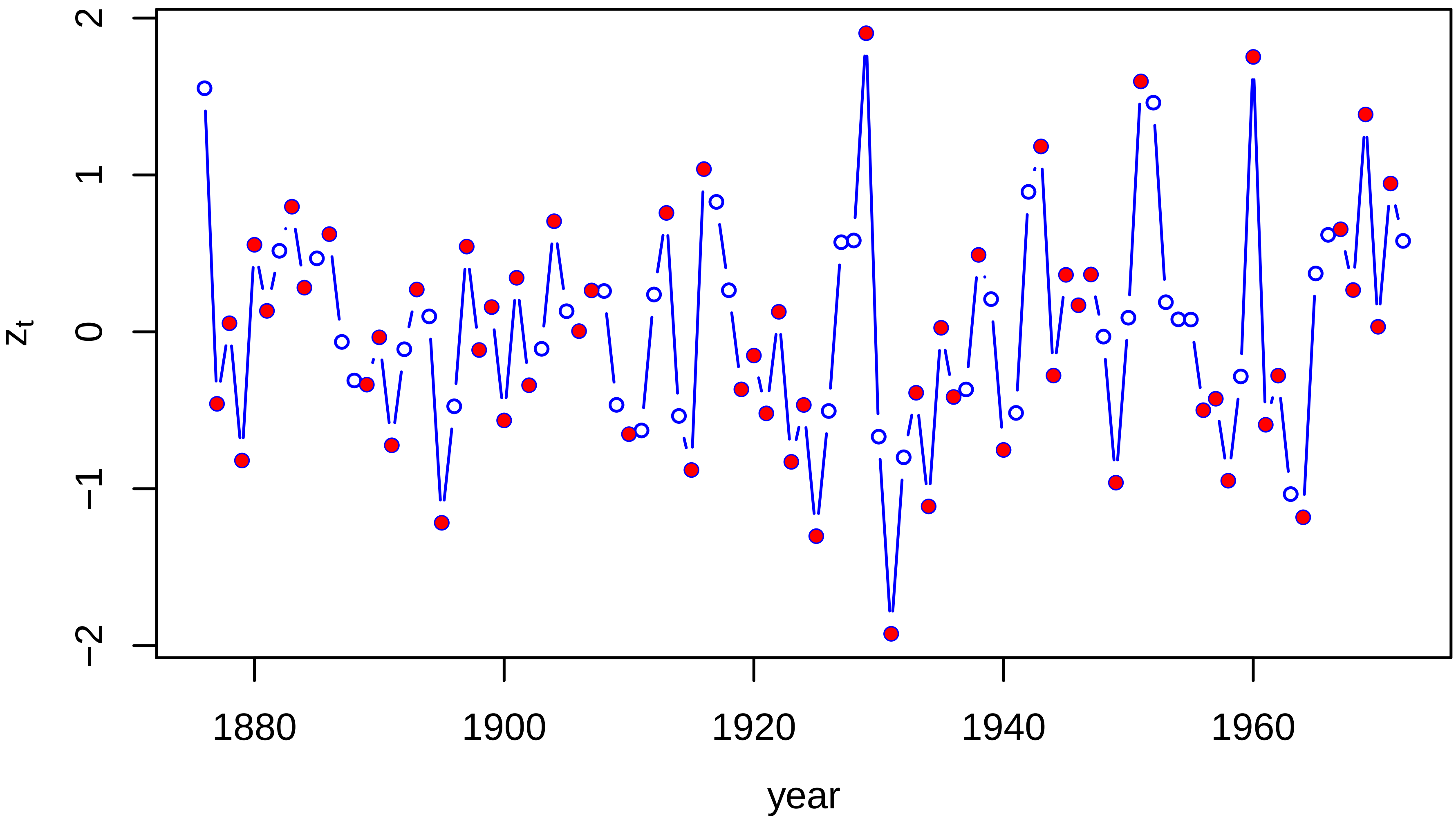
Turning Point Test: III

- note: cannot apply to all-star $\{x_t\}$ (has a discrete distribution)
- as example, $T = 60$ for Lake Huron $\{z_t\}$ (see next overhead)
- since $n = 97$, get $\mu_T \doteq 63.3$ and $\sigma_T^2 \doteq 16.9$ ($\sigma_T \doteq 4.1$), so

$$\frac{|T - \mu_T|}{\sigma_T} \doteq 0.81$$

- p-value is 0.42, so fail to reject null hypothesis

AR(1) Residuals $z_t = r_t - \hat{\phi} r_{t-1}$



Difference-Sign Test: I

- let S be the number of times $X_t > X_{t-1}$, $t = 2, \ldots, n$ (equivalent to number of times $X_t - X_{t-1}$ is positive)
- if $\{X_t\}$ is IID, then $\mu_S \stackrel{\text{def}}{=} E\{S\} = (n-1)/2$
- can shown that $\sigma_S^2 \stackrel{\text{def}}{=} \text{var}\,\{S\} = (n+1)/12$
- S is approximately equal in distribution to a $\mathcal{N}(\mu_S, \sigma_S^2)$ RV for large n
- large positive (negative) value of $S - \mu_S$ indicates presence of increasing (decreasing) trend (doesn't have power against alternative of cyclic variations)
- reject null hypothesis at level α if $|S - \mu_S|/\sigma_S > \Phi_{1-\alpha/2}$

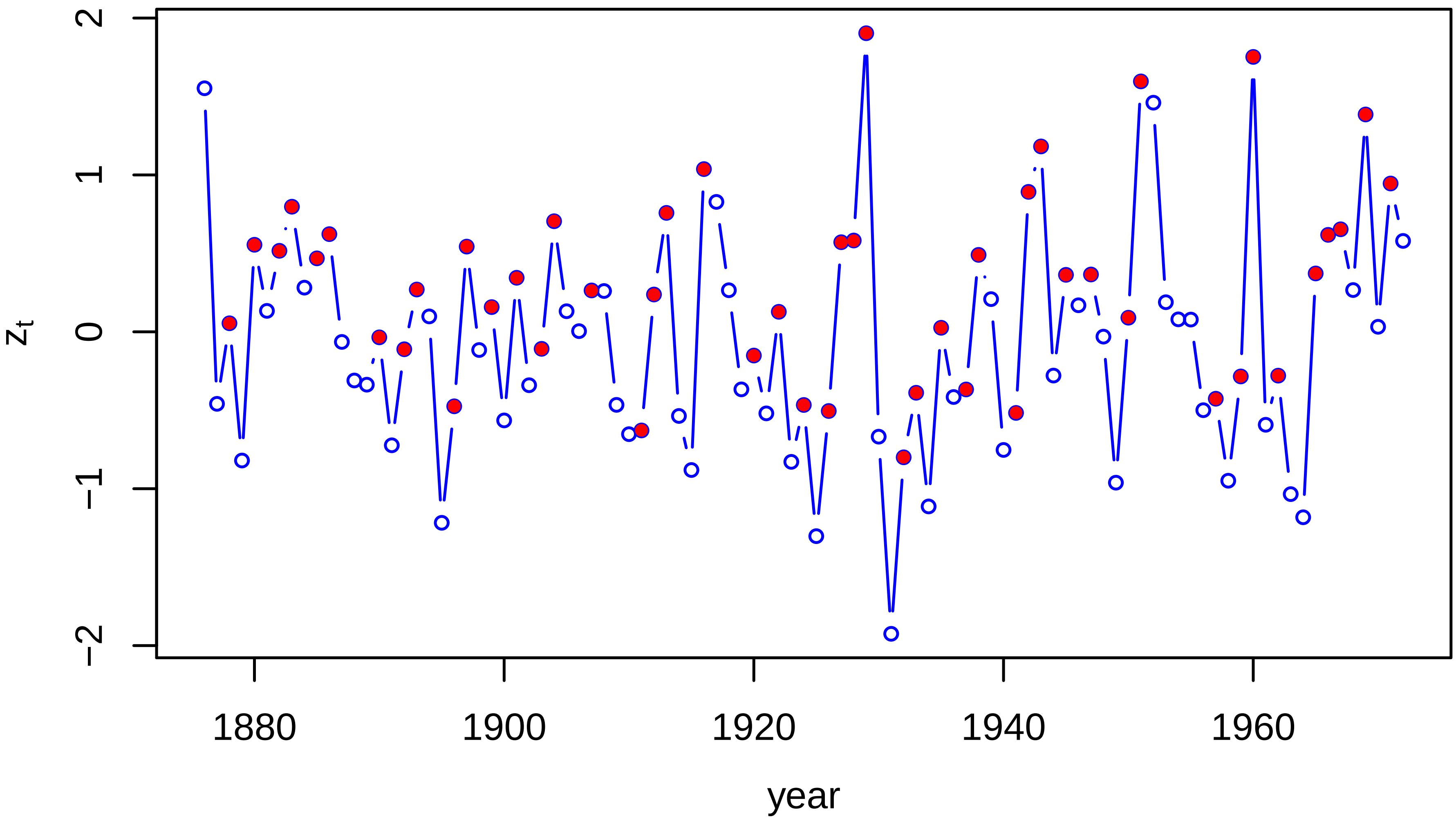
Difference-Sign Test: II

- note: cannot apply to all-star $\{x_t\}$ (has a discrete distribution)
- as example, $S = 48$ for Lake Huron $\{z_t\}$ (see next overhead)
- since $n = 97$, get $\mu_S = 48$ and $\sigma_S^2 \doteq 8.2$ ($\sigma_S \doteq 2.9$), so

$$\frac{|S - \mu_S|}{\sigma_S} = 0$$

- p-value is 1, so *totally* fail to reject null hypothesis!!!

AR(1) Residuals $z_t = r_t - \hat{\phi} r_{t-1}$



Rank Test: I

- given time series of length n, let P be the number of pairs (r, s) such that $X_r > X_s$ and $r > s$
- total number of pairs such that $r > s$ is
$$\binom{n}{2} = \frac{n(n-1)}{2}$$
- if $\{X_t\}$ is IID, each event of form $X_r > X_s$ has probability of $1/2$, so $\mu_P \stackrel{\text{def}}{=} E\{P\} = n(n-1)/4$
- can shown that $\sigma_P^2 \stackrel{\text{def}}{=} \text{var}\{P\} = n(n-1)(2n+5)/72$
- P is approximately equal in distribution to a $\mathcal{N}(\mu_P, \sigma_P^2)$ RV for large n
- large positive (negative) value of $P - \mu_P$ indicates presence of increasing (decreasing) trend

Rank Test: II

- reject null hypothesis at level α if $|P - \mu_P|/\sigma_P > \Phi_{1-\alpha/2}$
- note: cannot apply to all-star $\{x_t\}$ (has a discrete distribution)
- as example, $P = 2351$ for Lake Huron $\{z_t\}$
- since $n = 97$, get $\mu_P = 2328$ and $\sigma_P^2 \doteq 25737.3$ ($\sigma_P \doteq 160.4$), so
$$\frac{|P - \mu_P|}{\sigma_P} \doteq 0.14$$
- p-value is 0.89, so again fail to reject null hypothesis

Runs Test: I

- test designed for
  - binary-valued $\{x_t\}$ (will assume x_t is either -1 or 1)
  - continuous-valued $\{x_t\}$, with centering allowed; i.e., $\{x_t - c\}$, where, e.g., $c = \bar{x}$ is sample mean of x_t's
  - residuals $\{r_t\}$ (sample mean $\bar{r}$ should be close to 0)
- run is defined to be set of consecutive values in series, all of which are above or below zero
- as an example, here is pattern of American League and National League victories over 85 years in all-star $\{x_t\}$:

AAANANANAAANAAAANNNNANNAANNNNNNNNNNNNANNNNNNNNNNNANNANAAAAAANNNAAAAAAAAAAAANNNAAAAAAA

- above has a total of 25 runs (13 for AL, 12 for NL)

Runs Test: II

- runs test is based on number of runs in series, conditional on number of observed values above and below zero (test does *not* require underlying probability of value being greater than zero to be 50%)
- let n_+ and n_- denote number of, respectively, positive and negative values in series (thus $n_+ + n_- = n$)
- let n_r be number of runs
- under null hypothesis of IID, n_r is approximately normally distributed with mean μ and variance σ^2 given by

$$\mu = \frac{2n_+n_-}{n} + 1 \text{ and } \sigma^2 = \frac{(\mu - 1)(\mu - 2)}{n - 1}$$

Runs Test: III

- large positive (negative) value of $n_r - \mu$ indicates excessive choppiness (too much clumpiness) compared to what is reasonable under null hypothesis of IID
- reject null hypothesis at level α if $|n_r - \mu|/\sigma > \Phi_{1-\alpha/2}$
- using all-star $\{x_t\}$ as example, $n_r = 25$, $n_+ = 43$ & $n_- = 42$
- get $\mu \doteq 43.5$ and $\sigma^2 \doteq 21.0$ ($\sigma \doteq 4.6$), so

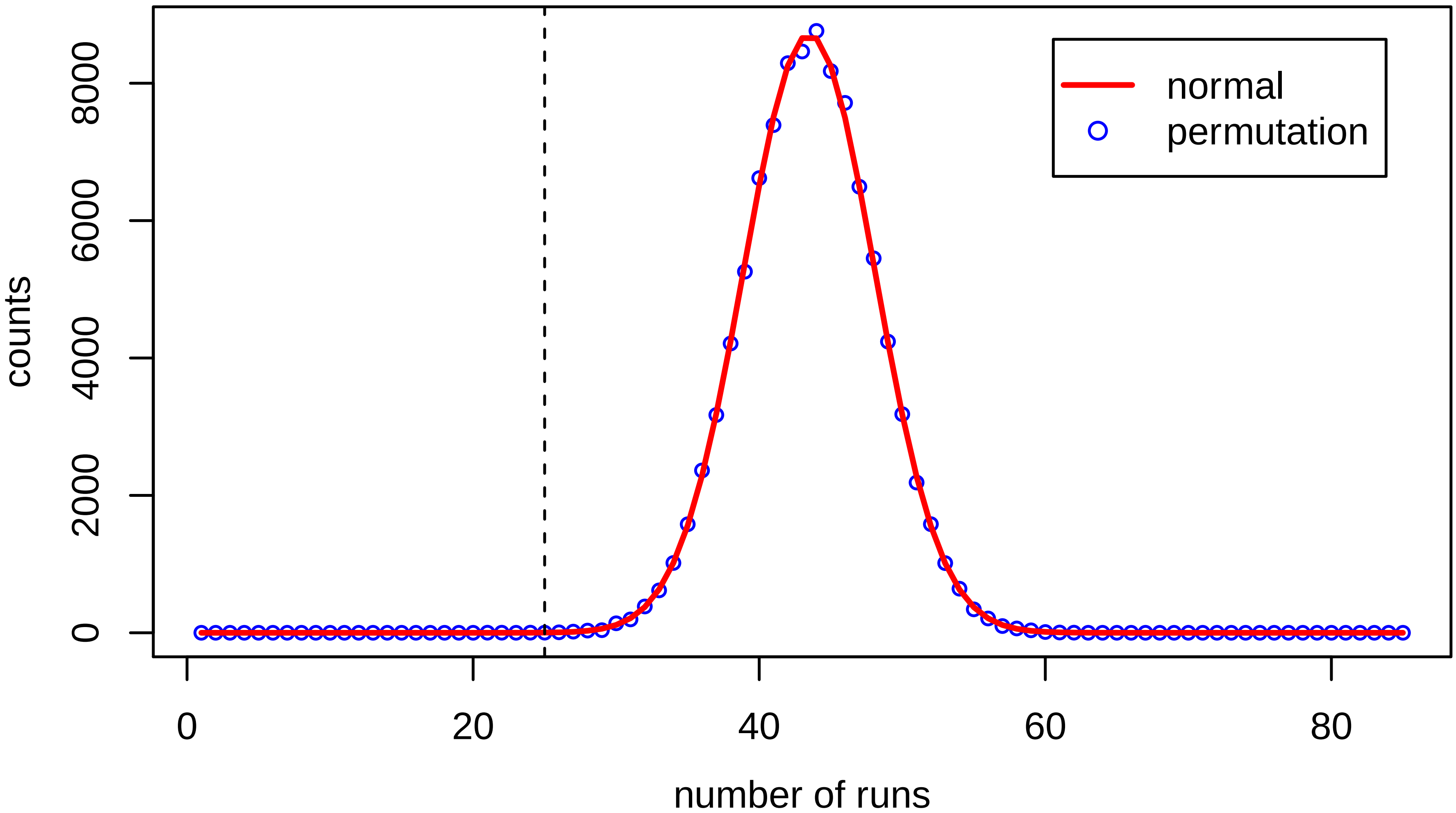
$$\frac{|n_r - \mu|}{\sigma} \doteq 4.0$$

- p-value is 0.00005, so null hypothesis is not tenable

Runs Test: IV

- exact distribution of n_r (conditional on n_+ & n_-) under null IID hypothesis follows from systematically permuting n values of series and determining number of runs in each permutation
- number of permutations is $n!$ – typically too large to enumerate, but can get good approximation to exact distribution by taking random samples of permutations (each sample is gotten by sampling n values from series *without* replacement)
- for all-star series, $n! = 85! \approx 3 \times 10^{128}$ – too large for enumeration, but can use, say, 100,000 randomly drawn permutations to get approximation to exact distribution of n_r
- following overhead compares this approximation to the one based on a normal distribution with $\mu \doteq 43.5$ and $\sigma^2 \doteq 21.0$ (vertical dashed line indicates 25, the actual number of runs)

Distribution of n_r under Null Hypothesis



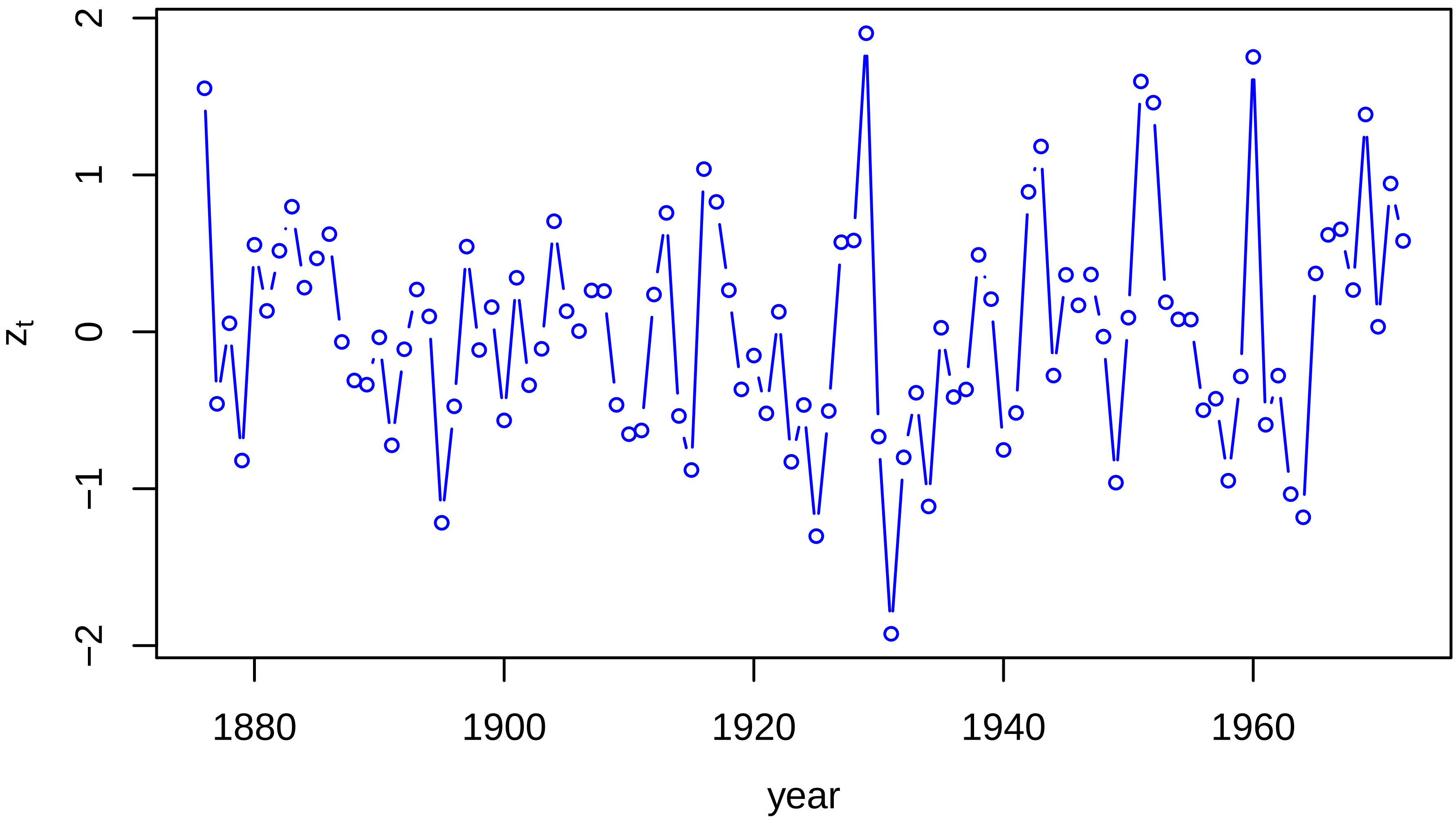
Runs Test: V

- for Lake Huron $\{z_t\}$, have $n_r = 37$, $n_+ = 51$ & $n_- = 46$
- get $\mu \doteq 49.4$ and $\sigma^2 \doteq 23.9$ ($\sigma \doteq 4.9$), so

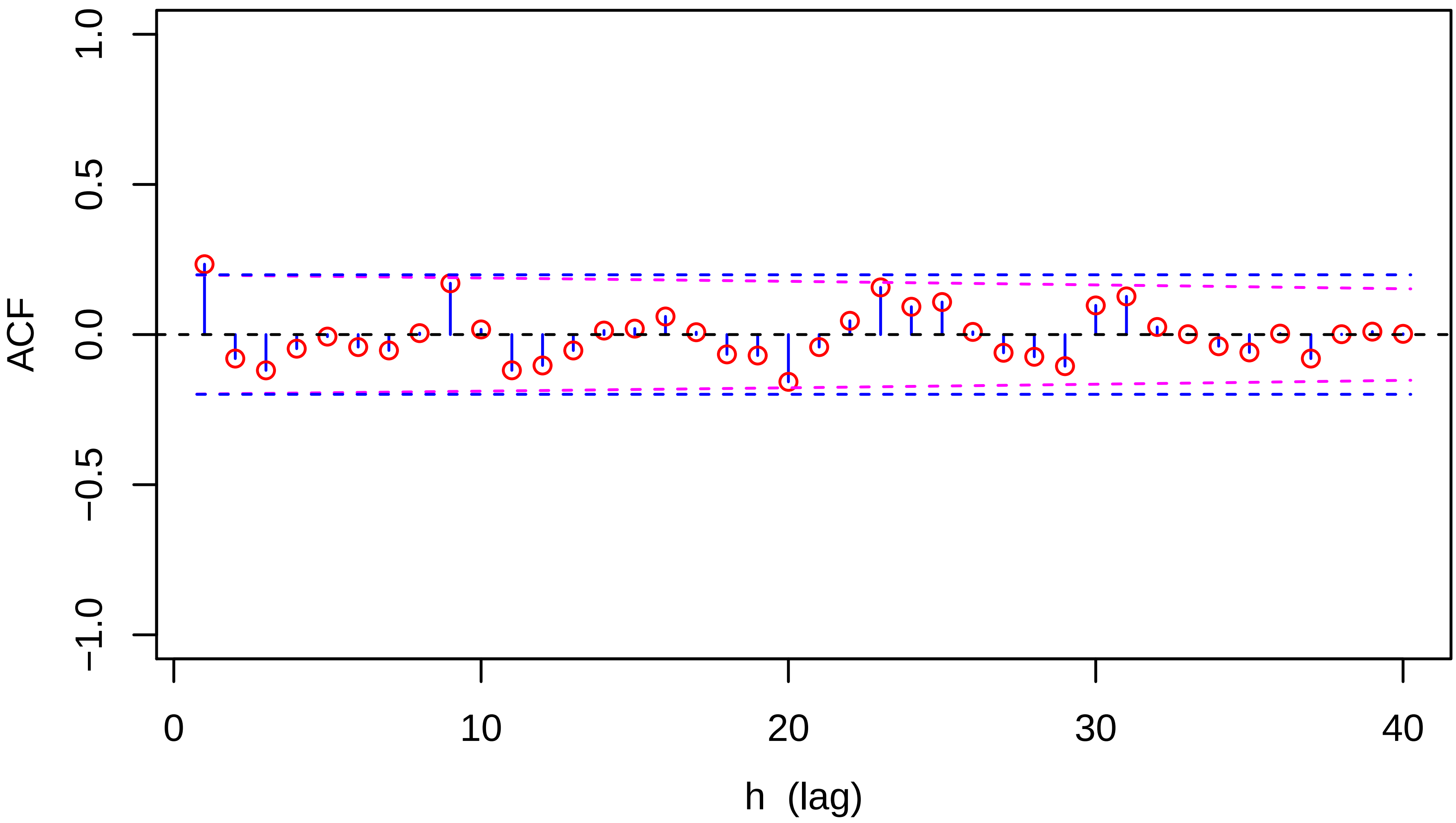
$$\frac{|n_r - \mu|}{\sigma} \doteq 2.5$$

- p-value is 0.011, so would reject null hypothesis at $\alpha = 0.05$ level of significance, but would fail to reject (just barely!) at more stringent $\alpha = 0.01$ level
- number of runs smaller than to be expected, suggesting clumpiness in data (agrees with positive correlation at unit lag)
- exercise: obtain p-value based on randomly drawn permutations

AR(1) Residuals $z_t = r_t - \hat{\phi} r_{t-1}$ (III–15)



Sample ACF for AR(1) Residuals z_t (III–17)



Testing for IID Noise/White Noise: III

- summary of IID tests for all-star $\{x_t\}$
  - informal test based on sample ACF: IID hypothesis untenable – both $\hat{\rho}_X(1)$ & $\hat{\rho}_X(2)$ well outside bounds reasonable for IID noise
  - portmanteau test: reject resoundingly for $h = 1, 2, \ldots, 10$
  - turning point, difference-sign & rank tests: inapplicable
  - runs test: reject (p-value is 0.00005)
- conclusion: IID hypothesis for all-star $\{x_t\}$ untenable
- action item: bet on American League in 2020!
  - note: Stat 519 last taught in 2018 – action item was to bet on American League, which would have been a winner and might have made some ex-Stat 519 students rich (if so, none have shared their wealth with their former instructor!)

Testing for IID Noise/White Noise: IV

- summary of IID tests for Lake Huron $\{z_t\}$
  - informal test based on sample ACF: IID hypothesis seems tenable, but $\hat{\rho}(1)$ outside of 95% CI
  - portmanteau test: reject at level $\alpha = 0.05$ for $h = 4, 5$ & 6 (p-values of 0.005, 0.02 & 0.046), but fail to reject for $h = 7, 8, \ldots, 20$ (p-values range from 0.07 to 0.44)
  - turning point test: fail to reject (p-value is 0.42)
  - difference-sign test: fail to reject (p-value is 1(!))
  - rank test: fail to reject (p-value is 0.89)
  - runs test: reject (p-value is 0.011)
- IID hypothesis for Lake Huron z_t's thus questionable since there is evidence for small – but significant – autocorrelation at least at lag $h = 1$

Testing for IID Noise/White Noise: V

- not clear how to combine results of tests together (multiple comparison problem)
- need to keep in mind various tests have differing strengths and weaknesses against particular alternative hypotheses
- three other tests:
  - informal test based on fitting AR models (will discuss later)
  - test based on correlation between sample ACF and sample partial autocorrelation function (PACF – will discuss later)
  - cumulative periodogram test (won't be discussed – test is covered under spectral analysis of time series)