THE TURNING POINT TEST FOR RESIDUALS WITH EQUALLY SPACED DESIGN POINTS

The classical "simple linear regression model" for y-on-x regression is to assume that $x_1, ..., x_n$ are non-random "design points," and

$$Y_i = a + bx_i + \varepsilon_i, \quad i = 1, ..., n,$$

where ε_i are i.i.d. $N(0, \sigma^2)$ for some unknown $\sigma > 0$. Doing the regression we get estimates $\widehat{a}, \widehat{b}$ of a and b and thus we get observed *residuals* defined by

$$\widehat{\varepsilon}_i = Y_i - \widehat{a} - \widehat{b}x_i, \ i = 1, ..., n.$$

Assume $n \geq 3$. For $i = 2, ..., n-1$, the residuals are said to have a turning point at i if either $\widehat{\varepsilon}_{i-1} < \widehat{\varepsilon}_i > \widehat{\varepsilon}_{i+1}$ (a local maximum at i) or $\widehat{\varepsilon}_{i-1} > \widehat{\varepsilon}_i < \widehat{\varepsilon}_{i+1}$ (a local minimum at i). Let $\widehat{T}_n$ be the observed number of turning points. There must be at least one, since $\widehat{\varepsilon}_i$ cannot be either monotone increasing or decreasing, if $\widehat{b}$ is computed correctly. So $1 \leq \widehat{T}_n \leq n - 2$.

The distribution of $\widehat{T}_n$ when the model is true depends on the design points. To get a unique distribution it's convenient to assume that the x_i are equally spaced: for some $h > 0$, $x_i - x_{i-1} = h$ for $i = 2, ..., n$.

In this handout, we're concerned with rejecting the simple linear regression hypothesis H_0 if $\widehat{T}_n$ is too small, and similarly also for quadratic regression.

The following tables are excerpted from the 2010 Ph. D. thesis of Xia Hua, "Testing regression models with residuals as data," available in MIT's DSpace. (The tables themselves are for 2-sided tests, allowing rejection if $\widehat{T}_n$ is too large.) For $n \leq 10$, Table 1 shows that H_0 can be rejected for $n = 7$ at the .05 level, and for $n = 8$ or 9 at the .01 level, if $\widehat{T}_n = 1$; and for $n = 9$ at the .05 level, or $n = 10$ at the .01 level, if $\widehat{T}_n \leq 2$.

Quoting from the thesis:

For the distribution of $\hat{T}_n$ under the null hypothesis, critical regions do not exist at the 0.05 or 0.01 levels for $n \leq 5$. The following tables give critical values of $\widehat{T}_n$ for $6 \leq n \leq 10$ and critical regions for tests at the 0.05 and 0.01 levels.

TABLE 1. Critical Values for $\widehat{T}_n$ of Simple Linear Regression at 0.05 and 0.01 Levels

n	$P_0(\widehat{T}_n \leq 1)$	$P_0(\widehat{T}_n \leq 2)$	$P_0(\widehat{T}_n \leq 3)$	$P_0(\widehat{T}_n = n-2)$	$\alpha = 0.05$	$\alpha = 0.01$
6	0.0705	0.3791	0.8180	0.1820	NA	NA
7	0.0200	0.1706	0.5410	0.1154	$\widehat{T}_n \leq 1$	NA
8	0.0049	0.0658	0.2988	0.0719	$\widehat{T}_n \leq 1$	$\widehat{T}_n \leq 1$
9**	0.0011	0.0223	0.1421	0.0456	$\widehat{T}_n \leq 2$	$\widehat{T}_n \leq 1$
10	0.0002	0.0067	0.0595	0.0287	$\widehat{T}_n \leq 2$ or $\widehat{T}_n = 8$*	$\widehat{T}_n \leq 2$

*$P_0(\widehat{T}_n = 0) = 0$ for any $n \geq 3$, $P_0(\widehat{T}_{10} = 7) = 0.1431$

**For $n = 9$, an alternative critical region is $\widehat{T}_n \leq 1$ or $\widehat{T}_n = 7$

RMD: The quadratic regression model, say H_Q, is that

$$Y_i = a + bx_i + cx_i^2 + \varepsilon_i$$

where again x_i are non-random design points and ε_i are i.i.d. $N(0,\sigma)^2$. The coefficients a, b, c can be estimated by least squares (with somewhat more algebra than for simple linear regression), and then we can define residuals as before. For quadratic regression, always $\widehat{T}_n \geq 2$. For $n \leq 10$, we can reject H_Q at level 0.05 for $n = 8$ or 9, and at level 0.01 for $n = 10$, if $\widehat{T}_n \leq 2$; for $n = 10$ we can reject H_Q at level .05 if $\widehat{T}_n \leq 3$ (footnote **) by the following table from the thesis:

Table 2. Critical Values for $\widehat{T}_n$ of Quadratic Regression at 0.05 and 0.01 Levels

n	$P_0(\widehat{T}_n \leq 2)$	$P_0(\widehat{T}_n \leq 3)$	$P_0(\widehat{T}_n = n-2)$	$\alpha = 0.05$	$\alpha = 0.01$
6	0.3247	0.7459	0.2541	NA	NA
7	0.1366	0.4540	0.1322	NA	NA
8	0.0497	0.2355	0.0878	$\widehat{T}_n \leq 2$	NA
9	0.0164	0.1070	0.0508	$\widehat{T}_n \leq 2$	NA
10**	0.0049	0.0433	0.0330	$\widehat{T}_n \leq 2$ or $\widehat{T}_n = 8$*	$\widehat{T}_n \leq 2$

*$P_0(\widehat{T}_n \leq 1) = 0$ for any $n \geq 3$, $P_0(\widehat{T}_{10} = 4) = 0.1695$, $P_0(\widehat{T}_{10} = 7) = 0.1431$

**For $n = 10$, an alternative critical region is $\widehat{T}_n \leq 3$

For $11 \leq n \leq 50$, let $\hat{k}_0$ be the largest k such that $P_0(\widehat{T}_n \leq k) \leq 0.025$ and $\hat{k}_1$ the smallest k such that $P_0(\widehat{T}_n \geq k) \leq 0.025$. Let $\hat{l}_0$ be the largest l such that $P_0(\widehat{T}_n \leq l) \leq 0.005$ and $\hat{l}_1$ the smallest l such that $P_0(\widehat{T}_n \geq l) \leq 0.005$. We define $\hat{F}_n^l(j) := P_0(\widehat{T}_n \leq j)$ and $\hat{F}_n^r(j) := P_0(\widehat{T}_n \geq j)$.

RMD: for rejection when $\widehat{T}_n$ is too small, we can ignore the right-hand halves of the following two tables. For each n, $11 \leq n \leq 50$, we can reject the hypothesis (linear or quadratic regression respectively in the two tables) at level .05 if $\widehat{T}_n \leq K_n$ where $K_n = \hat{k}_0$ or $\hat{k}_0 + 1$, depending on n as seen from the tables. Similarly we can reject at level 0.01 if $\widehat{T}_n \leq L_n$ where $L_n = \hat{l}_0$ or $\hat{l}_0 + 1$.

TABLE 3. Critical Values and Probabilities for $\widehat{T}_n$ of Simple Linear Regression

n	$\hat{l}_0$	$\hat{k}_0$	$\hat{F}^l_n(\hat{l}_0)$	$\hat{F}^l_n(\hat{l}_0+1)$	$\hat{F}^l_n(\hat{k}_0)$	$\hat{F}^l_n(\hat{k}_0+1)$	$\hat{k}_1$	$\hat{l}_1$	$\hat{F}^r_n(\hat{l}_1)$	$\hat{F}^r_n(\hat{l}_1-1)$	$\hat{F}^r_n(\hat{k}_1)$	$\hat{F}^r_n(\hat{k}_1-1)$
11	2	3	.0018	.0223	.0223	.1146	9	NA	NA	NA	.0183	.1203
12	2	3	.0005	.0077	.0077	.0505	10	NA	NA	NA	.0113	.0836
13	3	4	.0024	.0202	.0202	.0938	11	NA	NA	NA	.0075	.0579
14	3	4	.0007	.0075	.0075	.0429	12	12	.0045	.0396	.0045	.0396
15	4	5	.0026	.0180	.0180	.0765	13	13	.0030	.0271	.0030	.0271
16	4	5	.0008	.0070	.0070	.0358	13	14	.0019	.0185	.0185	.0837
17	5	6	.0025	.0156	.0156	.0629	14	15	.0011	.0124	.0124	.0605
18	5	6	.0009	.0064	.0064	.0301	15	16	.0008	.0085	.0085	.0436
19	6	7	.0024	.0134	.0134	.0516	16	17	.0005	.0057	.0057	.0311
20	6	7	.0009	.0057	.0057	.0252	16	17	.0038	.0220	.0220	.0798
21	7	8	.0023	.0115	.0115	.0427	17	18	.0025	.0155	.0155	.0594
22	8	9	.0049	.0210	.0210	.0668	18	19	.0016	.0108	.0108	.0439
23	8	9	.0020	.0098	.0098	.0352	19	20	.0011	.0076	.0076	.0323
24	9	10	.0043	.0175	.0175	.0550	19	21	.0008	.0236	.0236	.0745
25	9	10	.0018	.0083	.0083	.0292	20	21	.0038	.0172	.0172	.0568
26	10	11	.0037	.0147	.0147	.0455	21	22	.0026	.0124	.0124	.0428
27	10	12	.0015	.0069	.0241	.0669	22	23	.0021	.0092	.0092	.0324
28	11	12	.0033	.0124	.0124	.0378	22	24	.0011	.0238	.0238	.0687
29	11	13	.0014	.0202	.0202	.0556	23	24	.0045	.0177	.0177	.0531
30	12	13	.0028	.0103	.0103	.0312	24	25	.0031	.0130	.0130	.0407
31	13	14	.0050	.0168	.0168	.0461	25	26	.0023	.0096	.0096	.0310
32	13	14	.0024	.0087	.0087	.0260	25	27	.0017	.0071	.0236	.0632
33	14	15	.0043	.0141	.0141	.0384	26	27	.0050	.0176	.0176	.0492
34	14	16	.0021	.0073	.0217	.0546	27	28	.0037	.0133	.0133	.0383
35	15	16	.0036	.0117	.0117	.0320	28	29	.0027	.0099	.0099	.0296
36	15	17	.0017	.0061	.0181	.0456	28	30	.0021	.0075	.0228	.0579
37	16	17	.0030	.0097	.0097	.0266	29	31	.0015	.0056	.0175	.0458
38	16	18	.0015	.0051	.0151	.0381	30	31	.0040	.0131	.0131	.0357
39	17	19	.0025	.0081	.0222	.0526	31	32	.0032	.0102	.0102	.0281
40	18	19	.0043	.0126	.0126	.0318	31	33	.0022	.0075	.0216	.0528
41	18	20	.0022	.0069	.0187	.0443	32	34	.0013	.0053	.0164	.0417
42	19	20	.0036	.0105	.0105	.0266	33	34	.0040	.0126	.0126	.0330
43	19	21	.0018	.0057	.0156	.0372	34	35	.0030	.0097	.0097	.0260
44	20	22	.0030	.0088	.0224	.0504	34	36	.0023	.0074	.0204	.0482
45	21	22	.0048	.0131	.0131	.0312	35	37	.0018	.0057	.0159	.0386
46	21	23	.0026	.0074	.0188	.0425	36	37	.0043	.0123	.0123	.0307
47	22	23	.0040	.0109	.0109	.0263	36	38	.0034	.0097	.0245	.0546
48	22	24	.0021	.0061	.0157	.0358	37	39	.0026	.0075	.0194	.0442
49	23	25	.0034	.0092	.0222	.0478	38	40	.0018	.0056	.0150	.0354
50	23	25	.0018	.0052	.0133	.0303	39	40	.0041	.0116	.0116	.0282

*$P_0(\widehat{T}_{11} = 9) = 0.0182$, $P_0(\widehat{T}_{12} = 10) = 0.0115$, $P_0(\widehat{T}_{13} = 11) = 0.0073$, $P_0(\widehat{T}_{13} = 10) = 0.0504$

*Critical regions for n=11,12,13 at the 0.01 level are $\widehat{T}_{11} \leq 2$, $\widehat{T}_{12} \leq 3$ and $\widehat{T}_{13} \leq 3$ or $\widehat{T}_{13} = 11$.

TABLE 4. Critical Values and Probabilities for $\widehat{T}_n$ of Quadratic Regression

n	$\hat{l}_0$	$\hat{k}_0$	$\hat{F}_n^l(\hat{l}_0)$	$\hat{F}_n^l(\hat{l}_0+1)$	$\hat{F}_n^l(\hat{k}_0)$	$\hat{F}_n^l(\hat{k}_0+1)$	$\hat{k}_1$	$\hat{l}_1$	$\hat{F}_n^r(\hat{l}_1)$	$\hat{F}_n^r(\hat{l}_1-1)$	$\hat{F}_n^r(\hat{k}_1)$	$\hat{F}_n^r(\hat{k}_1-1)$
11	2	3	.0014	.0159	.0159	.1021	9	NA	NA	NA	.0199	.1328
12	2	3	.0003	.0052	.0052	.0443	10	NA	NA	NA	.0128	.0896
13	3	4	.0016	.0174	.0174	.0821	11	NA	NA	NA	.0078	.0619
14	3	4	.0004	.0063	.0063	.0371	12	NA	NA	NA	.0051	.0414
15	4	5	.0021	.0154	.0154	.0714	13	13	.0032	.0287	.0032	.0287
16	4	5	.0006	.0059	.0059	.0330	13	14	.0020	.0192	.0192	.0873
17	5	6	.0022	.0146	.0146	.0586	14	15	.0012	.0130	.0130	.0626
18	5	6	.0007	.0058	.0058	.0277	15	16	.0008	.0087	.0087	.0452
19	6	7	.0022	.0124	.0124	.0499	16	17	.0005	.0058	.0058	.0319
20	6	**8**	.0008	.0051	.0240	.0779	16	17	.0040	.0230	.0230	.0818
21	7	8	.0021	.0109	.0109	.0404	17	18	.0027	.0159	.0159	.0607
22	8	9	.0047	.0201	.0201	.0651	18	19	.0017	.0111	.0111	.0448
23	8	9	.0019	.0093	.0093	.0340	19	20	.0012	.0080	.0080	.0331
24	9	10	.0041	.0168	.0168	.0530	19	21	.0007	.0053	.0238	.0755
25	9	10	.0017	.0080	.0080	.0284	20	21	.0038	.0177	.0177	.0576
26	10	11	.0036	.0140	.0140	.0444	21	22	.0025	.0125	.0125	.0432
27	10	12	.0015	.0068	.0235	.0653	22	23	.0018	.0091	.0091	.0326
28	11	12	.0030	.0119	.0119	.0369	22	24	.0012	.0065	.0243	.0698
29	11	13	.0013	.0058	.0196	.0546	23	24	.0046	.0180	.0180	.0539
30	12	13	.0027	.0098	.0098	.0302	24	25	.0032	.0132	.0132	.0412
31	13	14	.0050	.0165	.0165	.0455	25	26	.0022	.0095	.0095	.0314
32	13	14	.0023	.0084	.0084	.0254	25	27	.0016	.0071	.0238	.0641
33	14	15	.0041	.0136	.0136	.0377	26	27	.0050	.0179	.0179	.0497
34	14	16	.0020	.0071	.0213	.0540	27	28	.0037	.0135	.0135	.0389
35	15	16	.0034	.0115	.0115	.0315	28	29	.0027	.0100	.0100	.0298
36	15	17	.0017	.0060	.0178	.0450	28	30	.0018	.0072	.0226	.0581
37	16	17	.0030	.0097	.0097	.0264	29	31	.0013	.0053	.0171	.0456
38	**17**	18	.0050	.0149	.0149	.0380	30	31	.0040	.0132	.0132	.0358
39	17	19	.0026	.0082	.0223	.0521	31	32	.0030	.0101	.0101	.0281
40	18	19	.0041	.0125	.0125	.0316	31	33	.0021	.0075	.0218	.0534
41	18	20	.0022	.0068	.0185	.0435	32	34	.0015	.0055	.0166	.0420
42	19	20	.0035	.0104	.0104	.0263	33	34	.0042	.0128	.0128	.0332
43	19	21	.0018	.0057	.0154	.0368	34	35	.0031	.0099	.0099	.0263
44	20	22	.0030	.0086	.0224	.0500	34	36	.0022	.0074	.0206	.0484
45	21	22	.0047	.0129	.0129	.0310	35	37	.0017	.0057	.0161	.0390
46	21	23	.0025	.0073	.0186	.0423	36	37	.0044	.0125	.0125	.0310
47	22	23	.0040	.0108	.0108	.0259	36	38	.0032	.0096	.0244	.0547
48	22	24	.0020	.0059	.0152	.0352	37	39	.0024	.0075	.0193	.0442
49	23	25	.0034	.0092	.0220	.0474	38	40	.0017	.0054	.0149	.0353
50	**24**	25	.0050	.0131	.0131	.0300	39	40	.0043	.0119	.0119	.0286

*$P_0(\widehat{T}_{11} = 9) = 0.0199$, $P_0(\widehat{T}_{12} = 10) = 0.0128$, $P_0(\widehat{T}_{13} = 11) = 0.0078$, $P_0(\widehat{T}_{13} = 10) = 0.0541$
*Critical regions for n=11,12,13 at the 0.01 level are $\widehat{T}_{11} \leq 2$, $\widehat{T}_{12} \leq 3$ and $\widehat{T}_{13} \leq 3$ or $\widehat{T}_{13} = 11$.
*A critical region for n=14 at the 0.01 level is $\widehat{T}_{14} \leq 3$ or $\widehat{T}_{14} \geq 12$. Alternatively, one can use $\widehat{T}_{14} \leq 4$.
Bold fonts indicate entries different for simple linear and quadratic regressions.