CLASSROOM NOTES. For this department we hope to obtain brief papers dealing with the subjects currently taught in *undergraduate* mathematics. Occasionally we shall also accept notes dealing with material commonly encountered in the first year of graduate study. Classroom notes should be relevant to some actual classroom situation that can be expected to exist at a good number of institutions.

MATHEMATICAL EDUCATION. There is now a growing and very healthy concern with mathematicians' role as teachers. This department welcomes reports of experiments in novel methods of teaching as well as discussions of all educational aspects of our profession.

PROBLEMS. This department attracts many readers and is an especially valuable feature of the MONTHLY. It will continue to seek and publish interesting elementary and advanced problems in pure and applied mathematics, both classical and modern.

REVIEWS. Through the telegraphic reviews the MONTHLY will continue to provide as complete a coverage as possible of the current textbooks. At present the MONTHLY is the only journal that performs this service for the mathematical community. Extended reviews will continue to appear, and we hope that the number of classroom reviews will increase.

R. P. BOAS, *Editor*

BREAKING RECORDS AND BREAKING BOARDS

NED GLICK

1. Introduction. When I take observations in chronological sequence, how often will the outstanding record value be surpassed? For example, suppose that I register the annual total inches of precipitation or, to take a less gloomy statistic, the total hours of bright sunshine in Vancouver, where I live: what is the probability that next year will be a new maximum?

Breakthroughs are less likely later than early in a sequence of observations. The first observation necessarily must be a "record high." But, prior to observing any values, I know that the second of two numbers in random sequence has equal probability of being smaller or larger than the first. Hence the probability is exactly 50% that a second, independent observation will be a new record high surpassing the initial record, assuming that there cannot be an exact tie (if measurement is arbitrarily precise).

From the same perspective, there is probability 1/3 that a third trial will be a new maximum, since the last of three repeated observations is equally likely to be smallest, middle, or largest.... Similarly, all 10 ranks are equally likely for the tenth observation; so maximum rank for the tenth observation has probability 1/10.

The theoretical expected or average number of record highs in a chronological sequence of n independent observations is the sum of these probabilities: $1+(1/2)+(1/3)+\cdots+(1/n)$.

This result surprises my acquaintances who have no background in probability theory. I have here the beginning of a conversational ploy to hold a person's interest when she or he asks what work I do and I say, "I am a statistician." Otherwise this reply may ruin a fine conversation, as in the dinner scene depicted by William Kruskal [**22**].

My ploy proceeds to verify predictions from the simple probability model with some real weather records. It happens that these weather records excellently illustrate the pathos and problems of work with statistical data.

It is just as interesting to note that the simple model does *not* fit record breaking in athletic competitions. Race times, jump heights, and throw distances have improved over several decades, while rainfall fluctuations from year to year are "random." In fact, the frequencies of record highs and lows can be used to infer whether observations indicate linear trend or random sequence.

After the weather and sports, I talk about traffic. Curiously, the simple model of random record values says something about how cars tend to bunch together behind a slow vehicle.

The same probability model applies in yet another context: a sequential strategy to find the weakest item in a sample of boards or beams, and hence to establish "tolerance limits" for lumber strengths. My interest in record highs and lows actually began in discussions at the Western Forest Products Laboratory of the Canadian Forestry Service. Breaking a random, but usually small, fraction of the available beams can accomplish the same purpose as 100% destructive laboratory testing. This conservation of material illustrates the economic spirit of experimental design.

So one example, the frequency of record highs or lows, ties together several "applied" aspects of a statistician's livelihood. More surprising, the intuitive idea that any record can be beaten also leads to mathematical proof that the harmonic sum $1+(1/2)+(1/3)+\cdots$ grows without bound, becoming bigger than any finite number.

Harmonic divergence is the simplest of many "well known" limit theorems and paradoxes related to record breaking. The mathematical sections comprising the second part of my paper review primarily those results in the theory of record values which do not depend on the particular distribution of the basic sequence of observations. This mathematics is mostly at the level of an undergraduate discrete probability course in the spirit of William Feller's famous text [**12**]. I avoid differentiation and integration as much as possible.

2. Weather records. Weather records in Canada are kept by the Meteorological Branch of the federal Department of Transport. Vancouver's weather records, however, are largely products of "the colourful careers of the two Shearman brothers": T. S. H. Shearman, who "was officially appointed by the government as weather observer for the city of Vancouver" in April, 1905, and E. B. Shearman, who replaced his brother from 1915 to 1948, for which service he received the British Empire Medal. This history I found in a Department of Transport 1964 Annual Meteorological Summary [**25**], from which most of my figures are taken.

Monthly, as well as annual inches of precipitation and hours of sunshine are shown in Data Set 1 and Data Set 2, respectively. Looking down the February column, for example, I find 5 record high years in the precipitation data (1900, 1901, 1902, 1918, 1961) and 5 in the sunshine data. Each sequence includes 65 observations; so $1+(1/2)+(1/3)+\cdots+(1/65)=4.76$ is the theoretical expected number of record highs.

The 12 monthly precipitation sequences actually give an empirical average of 4.17 highs per sequence. Record lows have the same distribution as highs, if the case of absolutely nil precipitation

Data Set 1.

Monthly & Annual Total Precipitation (Rain + One-Tenth Snow) in Vancouver, B.C., in Inches

Year	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	Annual	
1900	7.24	5.95	10.29	4.51	4.20	5.42	1.05	3.60	1.61	9.20	10.00	9.22	72.29	1
1901	11.28	6.31	3.04	5.29	4.38	5.01	83	.22	2.65	5.20	14.06	8.09	66.36	2
1902	6.68	10.17	7.45	3.11	4.40	1.97	2.37	1.15	3.39	4.72	10.33	9.55	65.29	3
1903	8.23	2.60	6.88	3.78	3.68	3.56	1.12	1.07	8.35	5.72	11.36	4.21	60.56	4
1906	9.66	6.03	2.37	1.04	3.58	3.04	.45	.83	8.87	7.60	8.25	7.33	59.05	5
1907	9.32	8.30	2.39	4.13	1.44	1.43	1.70	1.36	4.51	1.76	13.23	8.02	57.59	6
1908	7.60	6.31	7.14	2.61	4.11	1.86	1.59	1.15	1.46	6.68	13.69	8.41	62.61	7
1909	6.20	8.15	4.31	1.23	3.76	1.69	2.45	1.43	2.23	7.06	15.66	4.28	58.45	8
1910	11.19	5.01	2.91	3.60	2.15	1.98	.24	1.38	2.47	9.04	10.62	8.79	59.38	9
1911	6.11	3.37	3.05	1.96	5.39	2.09	.92	1.23	4.41	2.24	12.68	8.82	52.27	10
1912	8.47	6.25	.89	3.92	2.35	2.28	1.54	5.86	2.84	4.64	9.21	8.80	57.05	11
1913	9.62	4.38	5.38	2.53	4.33	3.81	2.02	.85	3.89	6.19	10.08	3.95	57.03	12
1914	10.56	4.87	3.33	3.28	.74	3.58	.42	.75	6.86	6.37	10.18	2.84	53.78	13
1915	7.13	4.42	4.18	3.04	3.42	1.07	.91	.36	.80	8.83	5.41	10.36	49.93	14
1916	5.96	7.40	14.55	4.07	1.41	1.34	5:25	.58	1.28	2.16	6.37	5.71	56.08	15
1917	9.33	5.87	5.61	8.20	1.69	5.40	.48	.93	3.30	3.49	5.23	11.72	61.25	16
1918	11.05	10.50	7.48	1.70	1.15	1.00	2.29	4.59	.30	7.56	7.02	8.07	62.71	17
1919	7.57	7.35	6.61	4.47	3.60	1.02	.15	1.15	1.16	3.14	11.25	9.74	57.21	18
1920	8.92	1.21	5.20	2.51	1.94	3.06	.67	2.91	10.37	8.34	7.14	11.01	63.28	19
1921	9.39	6.66	2.53	3.62	2.52	3.64	.32	2.84	5.03	10.08	8.99	5.56	61.18	20
1922	3.15	4.75	3.44	2.63	2.46	.17	.02	2.01	5.76	3.26	2.63	10.35	40.63	21
1923	8.73	4.06	3.86	2.14	2.84	2.07	.52	.73	2.97	2.56	6.13	15.88	52.49	22
1924	8.26	8.86	1.16	3.84	.31	.91	.71	1.88	5.81	6.56	5.87	8.35	52.52	23
1925	12.16	6.01	4.24	2.44	2.20	.38	.74	2.36	.44	3.00	4.05	13.05	51.07	24
1926	7.62	6.70	2.48	2.58	4.17	.78	.36	2.10	3.26	6.37	7.80	8.99	53.21	25
1927	9.08	4.73	6.80	1.88	5.12	1.16	.94	3.74	3.07	7.22	8.39	6.92	59.05	26
1928	9.04	1.87	7.01	4.29	2.22	1.93	.47	.20	1.35	7.38	5.36	5.34	46.46	27
1929	3.05	1.64	4.47	4.81	1.25	3.24	1.41	1.50	1.77	3.57	2.54	8.58	37.83	28
1930	2.75	9.42	3.33	4.72	2.86	2.18	.08	.07	2.65	7.50	3.00	5.22	43.78	29
1931	11.24	4.95	6.35	4.57	1.23	5.59	.44	.61	7.14	5.22	8.34	11.92	67.60	30
1932	9.18	6.65	9.15	4.84	1.34	2.08	5.32	2.01	2.62	5.66	10.17	7.37	66.39	31
1933	9.57	6.32	6.43	.53	4.33	2.03	1.76	1.12	5.89	7.94	5.45	12.85	64.22	32
1934	11.90	3.02	5.57	1.18	3.44	.69	1.86	1.24	2.77	4.99	9.56	12.27	58.49	33
1935	20.65	3.75	8.41	2.28	.55	1.06	1.79	1.36	2.57	6.70	4.65	8.46	62.23	34
1936	9.37	4.37	6.12	3.31	5.23	3.13	1.65	2.11	3.51	4.21	1.84	10.63	55.48	35
1937	2.42	8.62	3.64	7.16	3.14	6.14	.47	3.45	1.84	8.14	11.01	10.94	66.97	36
1938	6.59	4.68	4.65	2.93	1.86	.78	.66	.74	1.68	5.74	6.44	13.53	50.28	37
1939	11.91	5.94	3.83	1.71	3.40	3.03	2.34	.68	3.20	6.72	12.35	11.78	66.89	38
1940	5.65	8.11	7.39	4.57	2.85	.30	1.52	1.70	2.15	10.85	4.91	10.47	60.47	39
1941	8.04	6.24	3.82	2.59	5.73	2.19	.48	2.12	8.26	9.86	6.93	9.15	65.41	40
1942	3.56	2.82	4.25	3.88	1.89	5.28	3.36	.47	.79	6.20	6.11	9.59	48.20	41
1943	4.25	4.69	3.94	4.16	3.28	1.20	1.10	2.12	1.82	7.16	3.29	9.16	46.17	42
1944	7.85	4.63	3.41	4.15	1.90	.95	.49	.92	4.86	5.87	8.97	3.76	47.76	43
1945	9.33	7.45	8.54	4.23	2.64	.94	.91	1.27	3.21	7.10	10.07	6.45	62.14	44
1946	11.30	9.36	8.79	7.54	.43	4.87	1.63	.78	1.94	5.91	6.19	8.40	67.14	45
1947	10.68	7.02	6.52	5.62	1.74	2.08	2.34	.49	2.18	10.29	5.22	13.32	67.50	46
1948	3.76	10.31	2.67	3.59	6.05	11.82	2.16	4.03	2.83	4.69	14.57	9.59	66.07	47
1949	.84	7.58	5.05	2.22	1.73	1.76	2.75	1.46	1.45	6.83	11.89	7.62	51.18	48
1950	6.48	10.07	9.42	5.01	2.61	1.48	2.15	2.93	1.65	10.24	5.01	10.50	67.55	49
1951	11.10	10.28	6.40	2.82	3.50	.37	.01	.90	3.64	5.66	6.71	6.17	57.56	50
1952	8.23	5.55	5.99	2.72	2.23	3.92	.40	1.27	1.08	2.13	2.34	7.97	43.83	51
1953	14.08	4.79	4.93	2.38	2.31	2.26	1.29	1.83	4.76	5.26	10.44	11.28	65.61	52

1954	10.00	8.39	2.58	3.92	2.13	2.79	2.11	4.35	4.45	3.15	16.10	9.39	69.36	53
1955	5.09	4.17	7.41	4.32	3.22	3.03	2.97	.31	1.87	7.08	13.59	7.14	60.20	54
1956	7.10	6.47	7.72	1.12	.84	6.57	.89	2.52	5.65	13.69	13.86	13.71	70.14	55
1957	3.52	3.79	6.96	2.87	1.72	2.99	2.81	1.84	1.72	3.88	4.59	9.24	45.93	56
1958	13.48	6.75	3.51	3.60	1.33	1.51	Nil	2.35	3.10	5.93	9.03	9.44	60.03	57
1959	8.51	6.83	8.97	3.66	2.78	3.94	.97	.80	7.03	4.99	8.09	8.24	64.81	58
1960	7.43	7.04	5.49	3.13	5.64	2.05	.02	4.15	2.08	11.00	7.06	7.13	62.22	59
1961	13.28	15.26	6.04	2.98	3.96	1.42	1.34	3.56	1.94	8.82	7.70	10.04	76.34	60
1962	7.57	2.37	4.35	5.02	2.87	1.42	1.20	5.35	3.12	5.77	11.55	12.49	63.08	61
1963	1.76	7.53	4.29	4.09	1.81	2.24	3.08	.83	1.86	8.71	9.33	12.92	58.45	62
1964	11.62	3.82	6.44	3.27	2.75	2.58	3.56	2.33	7.41	3.38	10.17	6.71	64.04	63
1965	9.54	10.67	2.30	2.48	2.89	0.62	0.51	2.57	0.67	8.67	6.68	7.43	55.03	64
1966	9.28	4.53	4.87	1.75	2.89	2.18	3.36	1.66	3.17	7.43	9.47	15.34	65.93	65

DATA SET 2.

Monthly and Annual Total Hours of Bright Sunshine in Vancouver, B.C.

Year	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	Annual	
1909	55.3	56.4	143.7	252.4	230.3	251.8	224.4	262.8	294.9	99.2	58.1	51.4	1880.7	1
1910	45.5	91.5	117.7	181.2	276.5	186.5	312.8	224.1	181.2	103.6	49.9	31.0	1801.5	2
1911	14.4	84.8	154.0	248.2	174.9	217.7	272.7	225.1	142.2	140.0	41.6	48.3	1763.9	3
1912	44.2	81.4	233.8	113.8	215.4	229.0	198.1	186.6	185.9	110.7	19.4	26.0	1644.3	4
1913	43.3	97.1	125.6	149.5	178.2	186.6	270.2	228.0	173.2	110.6	62.1	27.6	1652.0	5
1914	38.7	46.4	104.1	155.8	270.2	250.5	313.5	280.7	79.4	101.6	39.0	67.1	1747.0	6
1915	57.7	44.1	131.8	203.0	163.7	266.7	261.7	221.9	141.3	84.6	70.2	36.4	1683.1	7
1916	65.9	84.8	86.4	141.5	216.9	222.9	145.5	284.4	190.5	150.9	84.7	23.5	1697.9	8
1917	53.9	56.7	153.5	90.4	220.6	188.5	358.7	348.2	153.8	104.5	50.9	23.1	1802.8	9
1918	45.3	95.8	103.5	257.5	253.0	318.8	297.4	232.4	236.2	77.1	48.4	58.4	2023.8	10
1919	50.6	58.2	128.2	147.2	257.5	250.1	341.8	265.5	202.9	122.7	60.7	73.5	1958.9	11
1920	44.8	148.4	108.0	196.1	249.0	206.5	308.7	308.6	138.2	91.4	77.8	20.0	1897.5	12
1921	31.6	85.9	137.2	180.1	275.0	137.4	317.2	221.1	172.1	123.7	42.9	57.5	1781.7	13
1922	77.4	98.5	116.5	143.7	243.2	257.0	267.5	193.2	163.4	106.7	71.7	23.3	1762.1	14
1923	52.3	75.2	163.4	194.7	181.4	228.6	289.1	291.7	226.4	142.8	49.9	30.0	1925.5	15
1924	24.6	59.2	157.8	168.0	250.3	238.0	273.9	238.0	217.8	118.7	54.3	62.4	1863.0	16
1925	49.4	49.7	137.3	168.7	261.0	254.3	327.7	240.7	205.2	116.0	59.7	10.7	1880.4	17
1926	27.4	56.1	187.7	220.0	186.4	288.6	300.3	236.9	221.7	93.8	60.5	30.5	1909.9	18
1927	39.9	96.3	128.8	164.5	199.9	219.6	305.1	178.4	124.2	78.3	28.8	56.5	1720.3	19
1928	42.6	102.7	116.2	168.4	278.4	195.7	285.1	258.1	199.2	94.3	45.0	38.9	1824.6	20
1929	50.5	109.7	119.3	169.3	249.0	187.9	322.2	304.9	225.2	113.7	61.7	25.1	1938.5	21
1930	100.7	85.8	169.5	142.7	254.2	219.0	320.6	305.6	192.0	125.9	54.1	40.7	2010.8	22
1931	28.5	80.2	109.8	215.2	274.7	187.9	381.2	334.1	118.5	124.9	96.8	23.2	1975.0	23
1932	65.4	97.0	94.6	147.4	246.3	297.0	195.2	237.7	235.3	115.1	37.8	63.0	1831.8	24
1933	36.8	86.1	117.8	223.4	146.1	198.3	336.3	302.3	123.7	112.4	42.9	30.4	1756.5	25
1934	56.7	118.8	136.5	237.5	236.2	301.0	236.2	270.2	155.3	126.5	34.2	43.8	1952.9	26
1935	44.8	94.3	84.8	246.9	277.5	194.7	247.3	257.1	213.2	109.0	59.1	26.9	1855.6	27
1936	54.6	70.5	121.6	168.5	174.0	212.9	313.3	295.0	181.4	130.1	52.3	36.6	1810.8	28
1937	87.1	69.7	93.7	79.0	221.1	201.6	292.6	203.0	164.6	121.2	37.1	34.0	1604.7	29
1938	30.7	78.0	110.5	175.9	313.1	292.1	314.6	253.9	161.6	132.7	68.4	52.6	1984.1	30
1939	20.8	75.9	116.2	166.5	206.7	143.9	278.7	320.6	193.6	108.7	36.8	30.1	1698.5	31
1940	42.9	63.4	109.0	167.2	242.6	329.2	236.5	266.1	180.1	73.9	48.6	37.6	1798.0	32
1941	38.7	93.0	161.4	184.6	191.8	164.2	326.0	231.2	119.2	80.1	71.2	55.9	1717.3	33
1942	51.2	75.7	101.6	120.6	140.0	192.7	236.4	291.0	175.1	116.3	61.3	44.8	1606.7	34
1943	52.8	104.3	107.7	149.5	177.1	234.7	251.5	203.4	201.0	126.0	42.1	43.2	1693.3	35
1944	50.9	80.7	155.7	121.4	213.0	201.2	259.1	179.9	167.0	108.0	38.5	42.2	1617.6	36
1945	51.9	75.5	77.6	129.9	200.2	191.5	283.3	252.1	154.0	124.0	33.5	32.8	1606.3	37
1946	38.2	42.0	95.2	88.4	287.1	135.7	232.9	267.7	144.9	122.6	66.5	17.8	1539.0	38

DATA SET 2 (continued).

Monthly and Annual Total Hours of Bright Sunshine in Vancouver, B.C.

Year	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	Annual	
1947	35.8	94.6	149.0	169.5	223.1	199.3	259.5	264.5	193.2	72.9	54.5	15.9	1731.8	39
1948	45.9	62.2	134.7	115.0	143.0	244.1	207.4	130.6	184.5	118.5	63.2	30.5	1479.6	40
1949	83.3	91.9	105.2	128.5	271.7	237.9	247.3	165.6	209.6	136.6	50.6	43.5	1771.7	41
1950	58.5	48.0	63.0	158.4	231.5	223.0	309.4	294.3	222.6	55.0	33.9	20.1	1747.7	42
1951	35.9	82.2	121.7	290.0	203.7	300.2	316.6	276.8	201.2	81.5	33.3	29.1	1972.2	43
1952	8.9	80.4	83.3	157.1	221.0	183.0	332.2	248.4	215.1	153.1	52.3	21.8	1756.6	44
1953	17.3	62.8	94.3	118.9	189.9	101.6	278.3	195.5	169.0	107.4	37.0	20.0	1392.0	45
1954	30.7	49.9	155.6	143.3	218.2	170.9	240.7	134.1	116.3	100.0	38.6	23.4	1421.7	46
1955	12.8	78.7	106.8	191.2	182.9	162.0	196.5	309.1	148.2	81.3	51.5	37.1	1558.1	47
1956	35.0	46.6	102.2	237.4	312.2	124.6	333.5	278.5	147.6	75.5	40.5	24.8	1758.4	48
1957	69.1	79.1	96.6	159.5	248.1	190.7	201.7	245.8	227.1	135.2	66.3	36.2	1755.4	49
1958	22.6	38.0	131.8	197.4	308.9	220.5	365.0	286.7	164.4	114.6	51.3	10.9	1912.1	50
1959	43.7	56.2	77.3	182.9	236.3	219.9	331.9	220.3	128.5	92.8	71.6	32.1	1693.5	51
1960	50.9	124.9	120.2	165.8	135.6	199.4	388.1	181.6	210.7	115.6	75.3	81.3	1849.4	52
1961	45.0	30.3	80.4	119.9	203.5	322.6	310.9	293.8	152.8	137.8	80.8	20.9	1798.7	53
1962	42.1	77.7	116.8	96.0	119.5	232.8	245.5	148.8	170.2	87.0	35.4	28.3	1400.1	54
1963	69.2	62.6	98.2	91.3	277.6	146.1	177.5	225.3	166.5	78.6	42.0	18.5	1453.4	55
1964	18.9	99.7	67.4	154.9	212.5	124.0	210.8	229.5	169.1	151.6	82.4	54.0	1574.8	56
1965	39	76	222	168	253	330	310	225	179	111	46	52	2010	57
1966	30	84	95	189	263	181	258	281	153	112	46	26	1718	58
1967	42	83	123	151	177	275	319	340	215	74.3	74	47	1920	59
1968	46	156	77	175	254	224	334	216.1	177	73.6	43	60	1836	60
1969	70	103	152	106	280	247	311	215.8	131	165	71	21	1873	61
1970	46	132	163	176	228	262	288	297	192.0	149	75	40	2048	62
1971	37	86	127	152	264	112	352	260	150	127	42	41	2013	63
1972	61	72	75	136	232	168	351	325	165	182	73	64	1904	64
1973	59	88	106	219	259	207	280	263	192.4	74.1	39	29	1815	65

SUMMARY OF RECORD BREAKTHROUGHS IN 65-YEAR SEQUENCES

Precipitation

	Jan.	Feb.	Mar.	Apr.	May	June	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	Average
Record Highs	4	5	2	3	6	4	5	2	6	4	4	5	4.17
Record Lows	10	3	4	4	6	7	8	4	4	4	7	5	5.50
Reverse Highs	7	3	5	8	$4\frac{1}{2}$	5	3	4	4	6	5	2	4.71
Reverse Lows	3	5	4	3	$7\frac{1}{2}$	5	4	5	4	5	5	6	4.71

Sunshine

	Jan.	Feb.	Mar.	Apr.	May	June	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	Average
Record Highs	4	5	3	3	4	5	6	4	1	7	5	4	4.25
Record Lows	4	6	7	5	7	5	3	5	4	6	4	7	5.25
Reverse Highs	6	3	5	3	5	4	4	4	6	2	5	3	4.17
Reverse Lows	6	4	4	6	5	4	5	6	7	3	5	5	5.00

can be excluded (an almost valid assumption in Vancouver); and these 12 monthly precipitation sequences give an average of 5.42 lows. Similar counts of "backward" highs and lows are obtained reading the 12 sequences in reverse, from most recent back to the oldest observations. (Since the May precipitation for 1965, recorded with two decimals, ties the figure for 1966, I treat May 1965 as 1/2 in the counts of reverse order highs and lows.) Considering highs and lows, chronological and reverse order, the 48 sequences give an overall empirical average of 4.77 record breaking values per sequence (compare to the theoretical expectation of 4.76).

All these counts can be repeated for the monthly sunshine sequences in Data Set 2.

The accompanying Summary of Record Breakthroughs shows all 96 counts (24 monthly data sets, considering highs and lows, forward and backward sequences). More than half of these 65-year-long sequences have exactly 4 or 5 breakthroughs.

People I ask usually guess that there will be more than 5 record values in 65 years. In predicting the number of record highs for a much longer sequence, say 1,000 or 1,000,000 observations, human intuition, even among mathematicians and statisticians, is definitely extravagant compared to the simple model's expectations of 7.49 and 14.39, respectively.

Over a long time, say a hundred or a thousand or ten thousand years, there may be shifts or cycles of climate; so the basic model of "interchangeable" weather years may not fit well. If climate trends are not negligible, then the actual probabilities for record breaking will differ from what the simple model predicts ..... In fact, the observed frequencies of record highs and lows can be used to infer whether or not data are a random sample.*

3. Tests of randomness. At a meeting of the Royal Statistical Society about twenty-five years ago, F. G. Foster and A. Stuart [**14**] pointed out that record low and record high annual rainfalls at Oxford were much more rare than record breaking performances (low times or high distances) in annual track and field competitions of the British Amateur Athletic Association.

This contrast is not surprising: athletic recruiting and training have intensified over the past century; but no one has done much about the weather. Although athletic performances do fluctuate, there is an average trend over decades for national competitors (and, therefore, winners) to run faster, jump higher, or throw farther; while weather fluctuations over a century are more intuitively random, without dramatic linear trend.

Of course it is possible for 100 random observations to be ordered so that the sequence has as many as 10 or 50 or 100 record highs. But detailed calculation [**6**] shows that the *probability* of 10 or more record highs in a 100-long random sequence is less than 5%. Therefore, in a situation where data are less familiar than rainfalls or race times, the mere finding of many record highs or lows suggests that the data are not a simple random sample; that is, an alternative hypothesis should be sought to fit the data better.

Foster and Stuart [**14**] gave formal procedures using the sum or the difference of record high and record low frequencies to fit or to test the hypothesis of randomness. Other statisticians have also considered such inference procedures [**3**, **4**, **14**, **15**].

---

* Returning to the Shearman brothers, I note that their locale of weather observation moved several times. Nonetheless, official records report precipitation for "Vancouver (City)" or "Port Meteorological Office" through 1966. Beyond that date there are reports for 10 separate city locations, plus the University of British Columbia and Vancouver Airport. Since rainfalls at nearby city locations often differ by more than an inch in a single month, I have not extended the precipitation sequences in Data Set 1 beyond 1966. Also I omit 1903 and 1904 because some months of precipitation data are missing from official records for those years.

For years since 1964 the monthly sunshine figures in Data Set 2 are taken from annual reports by the British Columbia provincial Department of Agriculture [**5**], because I could not find relevant annual publications by the federal Department of Transport. Although the provincial figures are derived from the same source, these monthly sunshine totals are rounded to the nearest whole hour, dropping one decimal. The provincial practice of first rounding and then adding the monthly figures may yield a different annual total than the federal total obtained by adding the monthly figures and then rounding-off. Moreover, other discrepancies which I noticed in comparing federal and provincial figures for 1960–1964 confirm that there are fairly frequent copying or typographic errors in one or the other publication (or both). Also the federal 1964 summary makes a substantial mistake in summing its own monthly sunshine figures to obtain its 1964 annual total. Apart from this addition error (which I noticed by accident) I have made no attempt in Data Sets 1 and 2 to find and correct the various errors carried over from my sources.

To count record highs and lows in the monthly sunshine sequences, I have broken several ties between rounded-off figures by finding more exact values from *daily* weather summaries.

... And there you have examples of the complications which arise in compiling real data, no matter how easy the job seems before you do it.

It can be shown that the number of forward and backward record highs, say R_n and R'_n, for any sequence have exactly the same *joint* probability distribution as the forward record highs and lows, say R_n and L_n [**4**]. And, as sample size $n \to \infty$, the counts of record highs and lows become approximately independent [**14**].

4. Car caravans in a one-lane tunnel. When traffic moving in one direction is confined to a single lane, a slow car is likely to be followed closely by a queue of vehicles whose drivers wish to go faster, but who cannot pass. If there is no exit from this lane, then more and more following vehicles will catch up and be added to the slow moving "platoon" or "caravan" ... until there happens to be a following vehicle travelling at a lower speed. This vehicle will not catch up, but will accumulate its own caravan.

Thus cars whose drivers all desire different speeds in fact will travel in caravans at *actual* speeds determined by record lows in the sequence of *desired* speeds.

Applying the simple probability model to a random sequence of drivers, the frequency of record lows corresponds to the *number of caravans* formed by n drivers. And the numbers of trials between successive record breaking low values correspond to the *lengths* of caravans.

Since caravans will be successively slower, separations between caravans will increase as time passes. G. F. Newell [**29**] used this reasoning to explain why cars near the exit of a long tunnel tend to travel faster and in smaller bunches more widely separated than cars in the tunnel near the entrance. This model of traffic flow also has been mentioned by other authors [**13**, **19**, **29**, **38**].

5. Sequential strategy for destructive testing. Many products fail under stress. For example, a wood beam breaks when sufficient perpendicular force is applied to it; an electronic component ceases to function in an environment of too high temperature; and a battery dies under the stress of time. But the precise breaking stress or failure point varies even among "identical" items.

Suppose that I can observe an item's exact failure point in a laboratory by gradually increasing stress (force, temperature, time, etc.). From such destructive testing of 100 items I could find all their failure points, say $X_1, X_2, \ldots, X_{100}$. But now suppose that I only need to find the *weakest* item in my sample: I only want the *minimum* value among failure stresses $X_1, X_2, \ldots, X_{100}$. Then I need not stress most of the items to their failure points.

The minimum failure stress among any batch of items can be determined sequentially. Test the first item until it fails, and record its failure stress X_1. Stop the next test (short of failure) if the second specimen survives this amount: so the second specimen's failure stress X_2 is determined exactly if $X_2 < X_1$; otherwise obtain only the "censored" information that $X_2 > X_1$, and hence $X_1 = \min(X_1, X_2)$. In either case proceed to the third specimen and stop the test if this item survives a stress equal to $\min(X_1, X_2)$: so X_3 is observed only if $X_3 < \min(X_1, X_2)$; but $\min(X_1, X_2, X_3)$ is always determined .... In general, the i^{th} item survives its stress test if $X_i > \min(X_1, \ldots, X_{i-1}) = \min(X_1, \ldots, X_i)$; or the test concludes with stress-to-failure if $X_i = \min(X_1, \ldots, X_i) < \min(X_1, \ldots, X_{i-1})$. In either case, the value $\min(X_1, \ldots, X_i)$ is known after the i^{th} trial.

The items destroyed in this sequential procedure are those with "record low" failure points. The frequency of such record lows fits the same probability model as the lows in a sequence of weather records. For a sample of n items, the expected number of items destroyed is $1 + (1/2) + (1/3) + \cdots + (1/n)$. This harmonic sum grows very slowly compared to sample size n. For example, the sum is only 5.19 when $n = 100$ and is only 7.49 when $n = 1000$ (see Table C).

The sequential strategy to find the minimum value generalizes easily to find the $2, 3, \ldots,$ or j smallest values among $X_1, X_2, \ldots, X_n$. To begin, test j items until they fail, at stresses $X_1, X_2, \ldots, X_j$. Thereafter stop the i^{th} trial if the item survives the j lowest failure stresses among all $i-1$ previous specimens. The probability of stress-to-failure for the i^{th} item ($i > j$) is the probability that it is among the j smallest of i independent observations from the same continuous distribution: all ranks are equally likely for X_i, so the desired probability is j/i. The expected number of items destroyed is the sum of failure probabilities over all trials:

$$\underbrace{1+1+\cdots+1}_{j\text{ terms}}+\frac{j}{j+1}+\frac{j}{j+2}+\cdots+\frac{j}{n}=j\left(1+\frac{1}{j+1}+\frac{1}{j+2}+\cdots+\frac{1}{n}\right)$$

$$\leqq j\left(1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n}\right).$$

If j is much less than the sample size n, then so is the expected number of failures. For example, to find the weakest 4 items in a sample of 1000, I expect to destroy only about 26; and to find the weakest 8 items I expect to destroy less than 50.

For some sorts of failure distributions, notably the exponential ([**13**] page 41), the minimum or the j smallest failure stresses from a sample can be used to estimate the theoretical mean of the failure distribution. Usually, however, the sample minimum is used to estimate a low percentile rather than the mean or median of the failure distribution.

6. Tolerance limits for failure distributions. A building code may prohibit use of a particular type of structural component unless it has probability at least .95 of surviving some severe stress x. In other words, the failure distribution's fifth percentile should satisfy $x_{.05}>x$. It is safer to under-estimate the percentile $x_{.05}$ than to over-estimate; so consider the lowest failure point observed in laboratory testing. In a very large sample it is highly probable that the breaking strength of the single weakest item is below the $x_{.05}$ value. It can be calculated, for example, that sample size $n=90$ is sufficiently large to assure that $x_{.05}>\min(X_1,X_2,\ldots,X_{90})$ with probability .99; so when this sample minimum exceeds x, the inference $x_{.05}>x$ has ".99 confidence."

The smallest value in a random sample of size $n\geqq 90$ also is called a level .99 *tolerance limit* for $x_{.05}$, the fifth percentile of the sampled distribution.

The tolerance limit interpretation of small order statistics is about thirty-five years old [**26, 41, 45**]. I encountered this subject when I joined a Statistics Committee of the Study Group on Wood Stresses under auspices of the Canadian Standards Association. (Vancouver's lumber export is important enough to be illustrated on the cover of the Annual Meteorological Summary [**25**] cited before.) The 1970 "Tentative Method for Evaluating Allowable Properties for Grades of Structural Lumber" [**1**] mentioned the sample minimum ($n=90$) as a level .99 tolerance limit, although it did not suggest the sequential strategy for finding it.

The 2$^{\text{nd}}$ smallest sample value is a level .99 tolerance limit for $x_{.05}$ when sample size $n\geqq 130$; and the 3$^{\text{rd}}$ smallest sample value is a level .99 tolerance limit when $n\geqq 165$. In general, the j^{th} smallest sample value is a level .99 tolerance limit for the fifth percentile when $n\geqq n_j$, where values n_j are given in Table A for $j=1,2,3,\ldots,15$. The greater the value of j, the less the variability of the corresponding tolerance limit and the less its conservative bias.

This table also gives values m_j, the mean number of failures (items destroyed) in sequential examination of n_j items to find the j weakest among them. Because the ratio m_j/n_j is about 1/10, I can afford a sample size about 10 times the number of items I can actually afford to break. For example, Table A shows that I must sample at least 228 items to use the 5$^{\text{th}}$ weakest as a level .99 tolerance limit for $x_{.05}$; but I expect to destroy only 23.63 of the 228 sample items.

Table A can be used in finding not only level .99, but also level .75, .90, or .95 tolerance limits for the fifth percentile. As the confidence level attached to the sample minimum is increased from .75 to 99, notice that sample size must be tripled, from 27 to 90; but the increase in expected failures, from 3.89 to 5.08, is slight. Hence, level .99 may cost little more than level .75 tolerance limits.

Table B similarly gives sample sizes n_j and expected failure numbers m_j for finding percentile $x_{.01}$ tolerance limits, at confidence level .75, .90, .95, or .99.

By now my writing has gone past the point I could ever carry a conversation. Often it is just as well to leave some air of mystery about how to calculate the minimum sample sizes in Tables A and B. But, of course, for a mathematician this computation is quite elementary.

TABLE A. *Non-parametric Tolerance Limits for Fifth Percentile,* $x_{.05}$

The j^{th} smallest order statistic from a sample of n items is a non-parametric tolerance limit for the sampled distribution's fifth percentile if sample size $n \geqq n_j$, where n_j is given in the following tables for confidence level $\gamma = .75, .90, .95$, or $.99$; that is, $X_{(j)}^{(n_j)} < x_{.05}$ with probability $\simeq \gamma$. The sequential strategy to find the j^{th} order statistic is expected to destroy m_j of the n_j items sampled.

$\gamma = .75$

j	n_j	m_j	m_j/n_j
1	27*	3.89	0.144
2	53	8.11	0.153
3	75*	11.45	0.153
4	101*	15.66	0.155
5	124*	20.59	0.166
6	147*	24.73	0.168
7	170	28.86	0.170
8	192*	32.96	0.172
9	215	37.09	0.173
10	237	41.18	0.174
11	259	45.28	0.175
12	281	49.37	0.176
13	303	53.46	0.176
14	325	57.55	0.177
15	346*	61.60	0.178

$\gamma = .90$

j	n_j	m_j	m_j/n_j
1	45	4.39	0.098
2	76*	8.83	0.116
3	105	12.46	0.119
4	132	17.52	0.133
5	158	21.80	0.138
6	183*	26.04	0.142
7	208*	30.27	0.146
8	233*	34.50	0.148
9	257*	38.69	0.151
10	281*	42.88	0.153
11	305*	47.07	0.154
12	329*	51.26	0.156
13	353	55.44	0.157
14	376*	59.59	0.158
15	399*	63.74	0.160

$\gamma = .95$

j	n_j	m_j	m_j/n_j
1	58*	4.65	0.080
2	93	9.22	0.099
3	124	13.70	0.111
4	153	18.11	0.118
5	180*	22.45	0.125
6	207*	26.77	0.129
7	234	31.09	0.133
8	260	35.38	0.136
9	285*	39.62	0.139
10	311	43.90	0.141
11	336	48.14	0.143
12	361	52.37	0.145
13	385*	56.57	0.147
14	409*	60.77	0.149
15	434	65.00	0.150

$\gamma = .99$

j	n_j	m_j	m_j/n_j
1	90	5.08	0.056
2	130	9.90	0.076
3	165	14.56	0.088
4	197*	19.12	0.097
5	228*	23.63	0.104
6	258*	28.09	0.109
7	287*	32.52	0.113
8	315*	36.91	0.117
9	343*	41.29	0.120
10	371	45.66	0.123
11	398	50.00	0.126
12	425	54.33	0.129
13	451	58.63	0.130
14	477*	62.92	0.132
15	503*	67.21	0.134

* Asterisk indicates sample size n_j such that $P\{X_{(j)}^{(n_j)} < x_{.05}\} < \gamma$, but the probability is as close as possible to γ (i.e., for sample size $n_j + 1$, the probability $P\{X_{(j)}^{(n_j)} < x_{.05}\} > \gamma$ and the excess over γ is of greater magnitude than the deficit for sample size n_j). Absence of an asterisk means sample size n_j is such that $P\{X_{(j)}^{(n_j)} < x_{.05}\} \geqq \gamma$ and the probability is as close as possible to γ.

For any proportion p, the p^{th} *percentile* of a continuous probability distribution F can be defined as the unique value x_p such that $F(x_p) = p$. In terms of the distribution's inverse function, the p^{th} percentile $x_p = F^{-1}(p)$. It is not difficult (see [8], page 7) to show that the j^{th} smallest order statistic $X_{(j)}^{(n)}$ from n sample observations will satisfy $X_{(j)}^{(n)} < x_p$ with probability

$$P\{X_{(j)}^{(n)} < x_p\} = \sum_{i=j}^{n} \binom{n}{i} [F(x_p)]^i [1 - F(x_p)]^{n-i} = \sum_{i=j}^{n} \binom{n}{i} p^i (1-p)^{n-i} = 1 - \sum_{i=0}^{j-1} \binom{n}{i} p^i (1-p)^{n-i}.$$

TABLE B. *Non-parametric Tolerance Limits for First Percentile*, $x_{.01}$

The j^{th} smallest order statistic from a sample of n items is a non-parametric tolerance limit for the sampled distribution's first percentile if sample size $n \geqq n_j$, where n_j is given in the following tables for confidence level $\gamma = .75, .90, .95$, or $.99$; that is $X_{(j)}^{(n_j)} < x_{.01}$ with probability $\simeq \gamma$. The sequential strategy to find the j^{th} order statistic is expected to destroy m_j of the n_j items sampled.

$\gamma = .75$

j	n_j	m_j	m_j/n_j
1	138	5.51	0.040
2	268*	11.34	0.042
3	391*	17.14	0.044
4	510	22.92	0.045
5	626*	28.67	0.046
6	741*	34.42	0.046
7	855	40.15	0.047
8	967*	45.87	0.047

$\gamma = .90$

j	n_j	m_j	m_j/n_j
1	229*	6.01	0.026
2	388	12.08	0.031
3	531	18.06	0.034
4	666*	23.98	0.036
5	797*	29.88	0.037
6	925*	35.75	0.039

$\gamma = .95$

j	n_j	m_j	m_j/n_j
1	298*	6.28	0.021
2	473	12.47	0.026
3	627*	10.56	0.030
4	773	24.58	0.032
5	913	30.56	0.033

$\gamma = .99$

j	n_j	m_j	m_j/n_j
1	458*	6.71	0.015
2	661	13.14	0.020
3	837*	19.42	0.023
4	1001	25.62	0.026
5	1157	31.74	0.027

* Asterisk indicates sample size n_j such that $P\{X_{(j)}^{(n_j)} < x_{.01}\} < \gamma$, but the probability is as close as possible to γ (i.e., for sample size $n_j + 1$, the probability $P\{X_{(j)}^{(n_j)} < x_{.01}\} > \gamma$ and the excess over γ is of greater magnitude than the deficit for sample size n_j). Absence of an asterisk means sample size n_j is such that $P\{X_{(j)}^{(n_j)} < x_{.01}\} \geqq \gamma$ and the probability is as close as possible to γ.

The subtracted sum is a binomial tail probability which can be evaluated using binomial tables (or a normal approximation) or a calculator. Since this sum vanishes as sample size $n \to \infty$, one can find, for any fraction γ, a sample size $n_j(\gamma)$ so large that the event $x_p > X_{(j)}^{(n)}$ has probability $\geqq \gamma$ for all $n \geqq n_j(\gamma)$.

Clearly this construction of tolerance limits is *distribution-free*: sample size $n_j(\gamma)$ does not depend on the form of the continuous distribution function $F(x) = P\{X_i < x\}$.

7. Persistence of record breaking and divergence of the harmonic series. No matter what the present precipitation record may be, it is certain that eventually there will occur a year with more rain. No matter how many record breaking years have been counted to date, there will be always one more. Such unit increments make the count of record breaking years arbitrarily large as the years of observation increase indefinitely.

These statements about rainfall interpret a specific mathematical theorem (and suggest its proof). From this point onward I emphasize formal limit results rather than applications of the theory of record values.

Formally, suppose that identically distributed random variables $X_1, X_2, \ldots, X_n$ are *exchangeable* (in particular, it suffices that the sequence of random variables be *independent*). The observation X_i is called a *record high* or *upper record value* or *ladder value** if X_i strictly exceeds all previous values in the sequence. For example, there are 3 record highs in the first 10 years of January sunshine figures

* A term used by Feller **[12]**.

(Data Set 2):

55.3, 45.5, 14.4, 44.2, 43.3, 38.7, 57.7, 65.9, 53.9, 45.3.

In this sequence the record highs are $X_1 = 55.3$, $X_7 = 57.7$, and $X_8 = 65.9$. Assume that exact ties have zero probability (arbitrarily precise measurement from a continuous distribution). Then X_i is a record high if and only if $X_i = \max(X_1, \ldots, X_i)$. Since all i ranks are equally likely for X_i, an upper record value (maximum rank) has probability $1/i$ at the i^{th} trial.

Let R_n denote the number of record highs among the first n observations. A formal version of the statement at the beginning of this section is the following theorem:

The count of record highs $R_n \to \infty$ with probability one as sample size $n \to \infty$.

To prove this proposition, consider an initial sequence of n_1 observations $X_1, X_2, \ldots, X_{n_1}$ and a further batch of n_2 observations $X_{n_1+1}, \ldots, X_{n_1+n_2}$. The probability that this additional batch contains no new record value is

$$P\{R_{n_1} = R_{n_1+n_2}\} = P\{\max(X_1, \ldots, X_{n_1}) = \max(X_1, \ldots, X_{n_1+n_2})\} = \frac{n_1}{n_1 + n_2},$$

the ratio of the initial sample size divided by the total. By taking batch size n_2 sufficiently large, I can make this probability as small as I wish. And I can repeat this procedure for many successive batches, choosing sizes $n_2, n_3, \ldots, n_r$ to satisfy

$$\frac{n_1}{n_1 + n_2} < \varepsilon / r$$

$$\frac{(n_1 + n_2)}{(n_1 + n_2) + n_3} < \varepsilon / r$$

$$\vdots$$

$$\frac{(n_1 + \cdots + n_{r-1})}{(n_1 + \cdots + n_{r-1}) + n_r} < \varepsilon / r$$

so that, for arbitrary $\varepsilon > 0$,

$$\sum_{k=1}^{r} P\{\text{no record value in } k^{\text{th}} \text{ batch}\} \leqq r(\varepsilon / r) = \varepsilon.$$

But the probability of obtaining at least r record highs in a sequence with length $n \geqq n_1 + n_2 + \cdots + n_r$ is then

$$\begin{aligned} P\{R_n \geqq r\} &\geqq P\{\text{at least one record high in each of } r \text{ batches}\} \\ &= 1 - P\{\text{at least one of } r \text{ batches has no record high}\} \\ &\geqq 1 - \sum_{k=1}^{r} P\{\text{no record high in } k^{\text{th}} \text{ batch}\} \\ &> 1 - \varepsilon. \end{aligned}$$

Since ε is arbitrarily small, conclude that $P\{R_n \geqq r\} \to 1$, for arbitrarily large integer r. That is, the number of record highs $R_n \to \infty$ in probability as sequence length $n \to \infty$. Because the sequence $R_1 \leqq R_2 \leqq \cdots$ is monotone, it follows that also $R_n \to \infty$ with probability one.

Viewed in the theory of recurrent events [12], record breaking is a *persistent* phenomenon.

That the expectation $E(R_n)\to\infty$ is immediate: for any integer r and $n \geqq n_1+n_2+\cdots+n_r$ as above,

$$E(R_n)=\sum_{k=1}^{n} kP\{R_n=k\}\geqq \sum_{k=r}^{n} kP\{R_n=k\}$$
$$\geqq rP\{R_n\geqq r\}>r(1-\varepsilon).$$

In fact, the preceding argument is an obvious modification of the classic harmonic divergence proof with batch sizes $n_k=2^{k-1}$ and $n=n_1+n_2+\cdots+n_r$. The n^{th} partial harmonic sum is:

$$\begin{aligned}1+\tfrac{1}{2}+\tfrac{1}{3}+\cdots+1/n &= 1+(\tfrac{1}{2}+\tfrac{1}{3})\\ &\quad+(\tfrac{1}{4}+\tfrac{1}{5}+\tfrac{1}{6}+\tfrac{1}{7})\\ &\quad+\cdots\\ &\quad+\underbrace{\left(\frac{1}{2^{r-1}}+\cdots+\frac{1}{2^r-1}\right)}_{2^{r-1}\text{ terms}}\\ &>1+2/4+4/8+\cdots+2^{r-1}/2^r=(r+1)/2.\end{aligned}$$

The other harmonic divergence demonstration which is popular in calculus texts shows $\Sigma_{i=1}^{n} i^{-1}\approx \int_1^n x^{-1}dx=\ln(n)$.

Several arguments in subsequent sections exploit these properties of the harmonic series.

8. The frequency of record breaking. Define binary variates to indicate the trials at which record highs occur in the original sequence:

$$Y_i=\begin{cases}1 & \text{if } X_i=\max(X_1,X_2,\ldots,X_i)\\ 0 & \text{otherwise,}\end{cases}$$

with expected value and variance given by

$$E(Y_i)=P\{Y_i=1\}=P\{X_i=\max(X_1,\ldots,X_i)\}=1/i,$$
$$V(Y_i)=E(Y_i^2)-[E(Y_i)]^2 \qquad\qquad =1/i-1/i^2.$$

Any distinct pair Y_i and Y_j (say $i<j$) are uncorrelated since

$$\begin{aligned}E(Y_iY_j)&=P\{Y_i=1 \quad\text{and}\quad Y_j=1\}\\ &=P\{X_i=\max(X_1,\ldots,X_i)\quad\text{and}\quad X_j=\max(X_1,\ldots,X_j)\}\\ &=P\{X_i=\max(X_1,\ldots,X_i)<\max(X_{i+1},\ldots,X_j)=X_j\}\\ &=P\{X_i=\max(X_1,\ldots,X_i)\}\\ &\qquad\times P\{\max(X_1,\ldots,X_i)<\max(X_{i+1},\ldots,X_j)\}\\ &\qquad\times P\{X_j=\max(X_{i+1},\ldots,X_j)\}\\ &=\frac{1}{i}\frac{j-i}{j}\frac{1}{j-i}=\frac{1}{ij}\\ &=P\{Y_i=1\}\;P\{Y_j=1\}=E(Y_i)E(Y_j).\end{aligned}$$

Thus the random number of record highs among $X_1,X_2,\ldots,X_n$ is $R_n=\Sigma_{i=1}^{n} Y_i$ with expectation and variance given by

$$E(R_n) = \sum_{i=1}^{n} E(Y_i) = \sum_{i=1}^{n} 1/i,$$

$$V(R_n) = \sum_{i=1}^{n} V(Y_i) = \sum_{i=1}^{n} 1/i - \sum_{i=1}^{n} 1/i^2.$$

Since

$$\sum_{i=1}^{n} \frac{1}{i} - \ln(n) \to \text{Euler's constant} = .5772\ldots \text{ and } \sum_{i=1}^{n} \frac{1}{i^2} \to \frac{\pi^2}{6} = 1.6449\ldots,$$

logarithm tables (or electronic calculators) easily give approximations for $E(R_n)$ and $V(R_n)$. Also both expressions are evaluated numerically in Table C for selected values of n: I find that the actual numbers surprise mathematicians as well as people who know little about logarithms.

TABLE C

Expectation, variance, and standard deviation of R_n, the number of record values in a random sequence of n independent and identically distributed observations

	$E(R_n) = \sum_{i=1}^{n} 1/i$	$V(R_n) = \sum_{i=1}^{n} 1/i - \sum_{i=1}^{n} 1/i^2$	
n	$E(R_n)$	$V(R_n)$	$\sqrt{V(R_n)}$
2	1.50	.25	.50
3	1.83	.47	.69
4	2.08	.66	.81
5	2.28	.82	.91
6	2.45	.96	.98
7	2.59	1.08	1.04
8	2.72	1.19	1.09
9	2.83	1.29	1.14
10	2.93	1.38	1.17
20	3.60	2.00	1.41
30	3.99	2.38	1.54
40	4.28	2.66	1.63
50	4.50	2.87	1.70
60	4.68	3.05	1.75
65	4.76	3.13	1.77
70	4.83	3.20	1.79
80	4.97	3.33	1.83
90	5.08	3.45	1.86
100	5.19	3.55	1.88
200	5.88	4.24	2.06
300	6.28	4.64	2.15
400	6.57	4.93	2.22
500	6.79	5.15	2.27
600	6.97	5.33	2.31
700	7.13	5.49	2.34
800	7.26	5.62	2.37
900	7.38	5.74	2.40
1000	7.49	5.84	2.42
1,000,000	14.39	12.75	3.57

The variance $V(R_n)$ gives bounds on the probability of destroying too many items in a sequential strategy for destructive testing. In particular, a one-sided Chebyshev inequality ([13] page 152) implies that

$$P\{R_n \geqq r\} \leqq \frac{V(R_n)}{V(R_n)+[r-E(R_n)]^2}.$$

For example, the values of $E(R_n)$ and $V(R_n)$ in Table C imply that $P\{R_{60} \geqq 9\} \leqq .14$ and $P\{R_{1000} \geqq 18\} \leqq .05$ (and actually these bounds are quite conservative since detailed calculation [6] shows $P\{R_{60} \geqq 9\} = .022$).

In a later section I show that the binary variates $Y_1, Y_2, Y_3, \ldots$ are *independent* as well as pairwise uncorrelated. Because of this independence, general limit theorems can be invoked to show more precisely how the random sum $R_n = \sum_{i=1}^n Y_i$ grows probabilistically like $\ln(n)$. I have noted that

(i) $$E(R_n) = \sum_{i=1}^n E(Y_i) = \sum_{i=1}^n \frac{1}{i} \rightarrow \infty$$

and that $\ln(n)$ approximates the expectation $E(R_n)$ for large n, so that

$$\frac{V(Y_n)}{[E(R_n)]^2} \approx \frac{1}{n[\ln(n)]^2}$$

and consequently

(ii) $$\sum_{n=1}^{\infty} \left\{ \frac{V(Y_n)}{\left[\sum_{i=1}^n E(Y_i)\right]^2} \right\} \approx \sum_{n=1}^{\infty} \frac{V(Y_n)}{[E(R_n)]^2} < \infty.$$

Invoking Kolmogorov's convergence criterion for sums of independent random variables ([24] page 238), conditions (i) and (ii) imply that the sequence $R_n/E(R_n) \rightarrow 1$; and hence Alfréd Renyi [30] obtained the following "strong law of large numbers" for the frequency of record highs (since $E(R_n)/\ln(n) \rightarrow 1$):

$$R_n/\ln(n) \rightarrow 1 \ \textit{with probability one as sample size } n \rightarrow \infty,$$

which implies the divergence $R_n \rightarrow \infty$ discussed in Section 7.

Similarly, the criterion of Liapounov ([24] page 275) gives a "central limit theorem" for the frequency of record highs: as sample size $n \rightarrow \infty$, the limit distribution of $(R_n - \ln(n))/\sqrt{\ln(n)}$ is normal with mean $= 0$ and variance $= 1$. (Renyi also gave a "law of the iterated logarithm." Resnick [33] later derived all these limit theorems from results concerning counts in a continuous-time Poisson process; see the end of Section 12, below.) But "the asymptotic distribution is a very poor approximation for $n \leqq 1000$ say, which is the only region of much statistical importance" [4].

The exact probability that $R_n = r$ is complicated. Instead, consider $R_{2n} - R_n$, the number of record highs over trials $n+1, n+2, \ldots, 2n$. In particular,

$$\begin{aligned} P\{R_{2n} - R_n = 0\} &= P\{\text{no record high for } n+1 \leqq i \leqq 2n\} \\ &= P\{\max(X_1, \ldots, X_n) = \max(X_1, \ldots, X_{2n})\} \\ &= n/2n = 1/2. \end{aligned}$$

In general, the count $R_{2n} - R_n$ has asymptotically a Poisson distribution with mean $= \ln(2)$: that is, for $k = 0, 1, 2, \ldots,$

$$P\{R_{2n} - R_n = k\} \rightarrow \frac{[\ln(2)]^k}{2k!} \quad \text{as } n \rightarrow \infty.$$

One proof uses the probability generating functions $g_i(s) = E(s^{Y_i}) = 1 + (s-1)/i$ associated with the binary variates Y_i. Since $R_{2n} - R_n = Y_{n+1} + Y_{n+2} + \cdots + Y_{2n}$, and the binary variates are independent, this sum has generating function given by

$$h_n(s) = \prod_{i=n+1}^{2n} g_i(s) = \prod_{i=n+1}^{2n} [1 + (s-1)/i].$$

Take logarithms of both sides and use the Taylor series expansion $\ln(1+x) = x - x^2/2 + x^3/3 - \cdots$ to obtain

$$\begin{aligned}\ln[h_n(s)] &= \sum_{i=n+1}^{2n} \ln[1 + (s-1)/i] \\ &= \sum_{i=n+1}^{2n} \left[\frac{s-1}{i} - \frac{(s-1)^2}{2i^2} + \frac{(s-1)^3}{3i^3} - \cdots\right] \\ &= (s-1)\sum_{i=n+1}^{2n} \frac{1}{i} - \frac{(s-1)^2}{2}\sum_{i=n+1}^{2n}\frac{1}{i^2} + \frac{(s-1)^3}{3}\sum_{i=n+1}^{2n}\frac{1}{i^3} - \cdots.\end{aligned}$$

As $n \to \infty$, all of the above series tails vanish except for the harmonic initial term, which converges to $(s-1)[\ln(2n) - \ln(n)] = (s-1)\ln(2)$. This limit is the logarithm of a Poisson generating function with parameter $\ln(2)$.... The argument can be generalized to prove the following theorem first stated by Dwass **[10]**:

As sample size $n \to \infty$, the frequency of record highs among observations indexed by $\boldsymbol{a}n < i \leqq \boldsymbol{b}n$ (for any $\boldsymbol{b} > \boldsymbol{a} > 0$) is asymptotically a Poisson count with mean $\ln(\boldsymbol{b}/\boldsymbol{a})$.

9. Serial numbers of record breaking trials. Rather than R_n, the random number of record highs in n trials, consider now the random serial number N_r of the trial at which the r^{th} record high occurs. The number N_r is called simply the r^{th} *record value time.* Since I count the initial observation as a record value, $R_1 = 1$ and $N_1 = 1$. In general $N_r \geqq r$ and, for $n \geqq r$,

$$P\{N_r \leqq n\} = P\{R_n \geqq r\}.$$

In particular, the record time N_2 has probability distribution, over the integers $i = 2, 3, \ldots$, given by

$$P\{N_2 = i\} = P\{X_1 = \max(X_1, \ldots, X_{i-1}) < X_i\} = \frac{1}{(i-1)i}.$$

Since this probability is strictly decreasing for $i = 2, 3, \ldots$, the *mode* $= 2$; but the *mean* is

$$E(N_2) = \sum_{i=2}^{\infty} iP\{N_2 = i\} = \sum_{i=2}^{\infty} \frac{1}{i-1} = \infty.$$

Since $N_2 < N_3 < \cdots$, it follows that $E(N_r) = \infty$ for all $r \geqq 2$. More surprising, an argument in the next section shows that $E(N_{r+1} - N_r) = \infty$.

To study joint distribution of N_r, N_{r+1} I must indicate why the binary variates $Y_1, Y_2, \ldots, Y_n$ defined in the preceding section are *independent* (a condition needed earlier to prove the Rényi and Dwass limit theorems). Consider the event that $Y_1, Y_2, \ldots, Y_n$ include exactly r ones, at trials $1 = i_1 < i_2 < \cdots < i_r \leqq n$. This event corresponds to

$$\begin{array}{ccccc}\max(X_1, \ldots, X_{i_2-1}) & < & \max(X_{i_2}, \ldots, X_{i_3-1}) & < \cdots < & \max(X_{i_r}, \ldots, X_n), \\ \| & & \| & & \| \\ X_1 & & X_{i_2} & & X_{i_r}\end{array}$$

which is the intersection of independent events of the form

$$\max(X_1,\ldots,X_{i-1}) < \max(X_i,\ldots,X_j) = X_i,$$

with $i = i_{k-1}$ and $j = i_k - 1$ or $j = n$. This component probability is

$$\frac{j-(i-1)}{j}\;\frac{1}{j-(i-1)} = \frac{1}{j};$$

so the desired joint probability is the product

$$\frac{1}{(i_2-1)(i_3-1)\cdots(i_r-1)n}.$$

It is easy to check that the product of marginal probabilities for $Y_2 = 0$, $Y_3 = 0,\ldots$, $Y_{i_2-1} = 0$, $Y_{i_2} = 1$, $Y_{i_2+1} = 0,\ldots$ gives the same expression:

$$\frac{1}{2}\,\frac{2}{3}\,\frac{3}{4}\cdots\frac{i_2-2}{i_2-1}\cdot\frac{1}{i_2}\,\frac{i_2}{i_2+1}\cdots = \frac{1}{(i_2-1)}\frac{1}{(i_3-1)}\cdots$$

The joint distribution of serial numbers $N_2, N_3,\ldots, N_r$ is therefore given by

$$P\{N_2 = i_2, N_3 = i_3,\ldots, N_r = i_r\} = \frac{1}{(i_2-1)(i_3-1)\cdots(i_r-1)i_r},$$

for $1 < i_2 < i_3 < \cdots < i_r$. The general marginal distribution is complicated, but expressions have been given by various authors [7, **19**, **30**, **37**]:

$$P\{N_r = n\} = |S_{n-1}^{r-1}|/n!,$$

where the numerator uses Stirling numbers of the first kind. Rényi [**30**] and David and Barton [**7**] used Stirling numbers to show that

$$\begin{aligned} P\{R_n = r\} &= \sum_{2\le i_2<i_3<\cdots<i_r\le n} P\{R_n = r;\, N_2 = i_2, N_3 = i_3,\ldots, N_r = i_r\} \\ &= \frac{1}{n}\sum_{2\le i_2<i_3<\cdots<i_r\le n}\frac{1}{(i_2-1)(i_3-1)\cdots(i_r-1)} \\ &= \frac{|S_n^{r-1}|}{(r-1)!} \approx \frac{[\ln(n)]^{r-1}}{n(r-1)!} \end{aligned}$$

for large sample size n. But this approximation also can be derived using generating functions as in the preceding section. (Or see [**21**], page 267.)

Rényi [**30**] treated limit theorems for N_r as duals of theorems for R_n. First, there is a "strong law of large numbers":

$$\frac{\ln(N_r)}{r} \to 1 \textit{ with probability one as } r\to\infty;$$

or, equivalently,

$$N_r^{1/r} \to e \textit{ with probability one.}$$

This last result says that almost every sequence of chance observations $X_1, X_2, X_3,\ldots$ from a continuous distribution will give an arbitrarily precise estimate of the mathematical constant e, even when the distribution function is unknown to me! I need only the serial numbers of trials at which record highs occur. (This approach differs somewhat from statistical determinations of e or of π in experiments such as "matching" or "Buffon's needle problem" [**12**]. In those situations, a particular event probability is a function of e or of π; and this probability is estimated by the relative frequency in repeated trials.)

Also Rényi [30] stated a "central limit theorem" for the random variables N_r: as $r \to \infty$, the distribution of $(\ln(N_r) - r)/\sqrt{r}$ is asymptotically normal with mean = 0 and variance = 1.

Rényi also gave a "law of the iterated logarithm" for N_r. Resnick [33] again derived Rényi's limit theorems for record times from a more sophisticated result. And Vervaat ([44] page 323) gave a Wiener process generalization of Rényi's N_r central limit theorem.

Returning to the distribution of the second record time N_2, described at the beginning of this section, notice that for $n = 2, 3, 4, \dots$,

$$P\left\{\frac{1}{N_2} < \frac{1}{n}\right\} = P\{N_2 > n\} = \sum_{i=n+1}^{\infty} P\{N_2 = i\}$$

$$= \sum_{i=n+1}^{\infty} \frac{1}{(i-1)i} = \sum_{i=n+1}^{\infty} \left(\frac{1}{i-1} - \frac{1}{i}\right) = \frac{1}{n}.$$

A similar result holds for record time N_3 and *conditional* probability of the event $N_3 > n$, given that the preceding record time $N_2 = m$. More generally, for $n > m \geqq r$,

$$\begin{aligned} P\left\{\frac{N_r}{N_{r+1}} < \frac{m}{n} \,\middle|\, N_r = m\right\} &= P\{N_{r+1} > n \mid N_r = m\} \\ &= P\{\text{no record high for } m+1 \leqq i \leqq n \mid N_r = m\} \\ &= P\{\max(X_1, \dots, X_m) = \max(X_1, \dots, X_n) \mid N_r = m\} \\ &= m/n. \end{aligned}$$

Any real x in the interval $(0, 1)$ can be approximated by a rational ratio m/n; and Tata [43] gave the following limit theorem:

The distribution of the ratio N_r/N_{r+1} *is asymptotically uniform over the unit interval: that is, for* $0 < x < 1$,

$$P\left\{\frac{N_r}{N_{r+1}} < x\right\} \to x \quad \textit{as } r \to \infty.$$

The proof suggests duality between this result and the theorem of Dwass in the preceding section. Given that the r^{th} record high occurs at trial N_r, the conditional probability

$$\begin{aligned} P\left\{\frac{N_r}{N_{r+1}} < \frac{1}{2} \,\middle|\, N_r\right\} &= P\{N_{r+1} > 2N_r \mid N_r\} \\ &= P\{\text{no record high for } N_r + 1 \leqq i \leqq 2N_r \mid N_r\} \\ &= P\{R_{2N_r} - R_{N_r} = 0 \mid N_r\} = N_r/2N_r \\ &= 1/2, \end{aligned}$$

independent of N_r. Similarly, for $0 < x < 1$,

$$P\left\{\frac{N_r}{N_{r+1}} < x \,\middle|\, N_r\right\} = P\left\{N_{r+1} > \frac{N_r}{x} \,\middle|\, N_r\right\} = \frac{N_r}{[N_r/x]} \to x$$

as $r \to \infty$, independent of N_r. (Here $[N_r/x]$ means "the largest integer $\leqq N_r/x$.")

The argument also indicates, as Shorrock [37] and later Resnick [33] proved, that successive ratios $N_r/N_{r+1}, N_{r+1}/N_{r+2}, \dots$ are *asymptotically independent* uniform variates. In other words, Shorrock's theorem shows that an *unknown* continuous distribution can be used to approximate a *uniform* random number generator! I need only the numbers of the trials at which record highs occur in the original randomly sampled $X_1, X_2, X_3, \dots$.

Tata's argument requires the asymptotic probability that the count $(R_{[n/x]}-R_n)=0$. In the preceding section, however, the theorem of Dwass gives more: the asymptotic Poisson probability that $(R_{[n/x]}-R_n)=k$, for $k=0,1,2,\ldots$. Therefore, for any positive integer s and real x in the unit interval,

$$P\left\{\frac{N_r}{N_{r+s}}<x \,\middle|\, N_r\right\}=P\left\{N_{r+s}>\frac{N_r}{x}\,\middle|\, N_r\right\}=P\{R_{[N_r/x]}-R_{N_r}<s \mid N_r\}$$

$$=\sum_{k=0}^{s-1} P\{R_{[N_r/x]}-R_{N_r}=k \mid N_r\}\rightarrow x\sum_{k=0}^{s-1}(-\ln(x))^k/k!$$

as $r\rightarrow\infty$, independent of N_r. Thus a dual of the Dwass theorem generalizes Tata's result: for $s=1,2,3,\ldots$ and $0<x<1$,

$$P\left\{\frac{N_r}{N_{r+s}}<x\right\}\rightarrow x\sum_{k=0}^{s-1}(-\ln(x))^k/k! \quad \text{as } r\rightarrow\infty.$$

10. Waiting times between record breaking trials. Now consider the wait from the r^{th} to the $(r+1)^{\text{st}}$ record breaking trial: denote this *inter-record waiting time* by $W_r=N_{r+1}-N_r$.

Record breaking must occur infinitely often; but it turns out that the inter-record waiting time distributions have mean $=\infty$, although mode $=1$. These results, first proved by Chandler [**6**], can be deduced from independence of the binary variates $Y_1, Y_2,\ldots$ via the following arguments.

$$\begin{aligned}P\{W_r=k \mid N_r\}&=P\{N_{r+1}=N_r+k \mid N_r\}\\&=P\{Y_i=0 \quad \text{for } N_r+1\leqq i\leqq N_r+k-1,\ Y_{N_r+k}=1 \mid N_r\}\\&=\frac{N_r}{N_r+1}\,\frac{N_r+1}{N_r+2}\cdots\frac{N_r+k-2}{N_r+k-1}\,\frac{1}{N_r+k}=\frac{N_r}{(N_r+k-1)(N_r+k)}.\end{aligned}$$

This expression is monotone decreasing as k increases; so, for any value of N_r, $P\{W_r=1 \mid N_r\}\geqq P\{W_r=k \mid N_r\}$, for $k=1,2,3,\ldots$. Taking expectations, $P\{W_r=1\}\geqq P\{W_r=k\}$; that is, the most probable value or mode $=1$.

$$\begin{aligned}P\{N_r=i,\ W_r=k\}&=P\{N_r=i,\ N_{r+1}=i+k\}\\&\geqq P\{N_2=2, N_3=3,\ldots, N_{r-1}=r-1, N_r=i, N_{r+1}=i+k\}\\&=\frac{1}{(2)(3)\cdots(r-2)(i-1)(i+k-1)(i+k)}\\&\geqq\frac{1}{r!(i-1)}\left(\frac{1}{k+i-1}-\frac{1}{k+i}\right)\end{aligned}$$

and hence, for any integers m and $i\geqq r$,

$$\begin{aligned}P\{N_r=i,\ W_r\geqq m\}&\geqq\frac{1}{r!(i-1)}\sum_{k=m}^{\infty}\left(\frac{1}{k+i-1}-\frac{1}{k+i}\right)\\&=\frac{1}{r!(i-1)(i-1+m)}=\frac{1}{r!m}\left(\frac{1}{i-1}-\frac{1}{i-1+m}\right).\end{aligned}$$

The marginal distribution of W_r satisfies

$$\begin{aligned}P\{W_r\geqq m\}&=\sum_{i=r}^{\infty}P\{N_r=i,\ W_r\geqq m\}\geqq\frac{1}{r!m}\sum_{i=r}^{\infty}\left(\frac{1}{i-1}-\frac{1}{i-1+m}\right)\\&\geqq\frac{1}{r!m}\sum_{i=r}^{r+m-1}\frac{1}{i-1};\end{aligned}$$

and therefore, for arbitrarily large integers m,

$$E(W_r)=\sum_{k=1}^{\infty} kP\{W_r=k\}\geqq mP\{W_r\geqq m\}$$

$$\geqq \frac{1}{r!}\sum_{i=r}^{r+m-1}\frac{1}{i-1}.$$

It follows from the divergence of the harmonic series that $E(W_r)=\infty$.

In other words, if I pay a fixed fee and receive in return a variable dollar amount equal to the waiting time W_r, then my expected gain is positive and my gamble is not a "fair game" no matter how high the fee.

An alternate demonstration that $E(W_r)=\infty$ assumes that observations $X_1, X_2, X_3, \ldots$ have exponential distribution $F(x)=P\{X_i<x\}=1-e^{-x}$, for $x>0$. This assumption is harmless because the preceding section shows that record times N_r, and hence inter-record waiting times W_r, have distributions which do not depend on F (only that it be continuous). The *conditional* probability that waiting time $W_r>m$, given that the r^{th} record value $X_{N_r}=x$, is just the probability that independent $X_i<x$ for indices $N_r+1\leqq i\leqq N_r+m$; viz., conditional probability is $[F(x)]^m=(1-e^{-x})^m$ for exponential sampling. Moreover, Section 11 shows that the record value X_{N_r} from exponential sampling has a gamma density. Hence the *unconditional* probability that $W_r>m$ can be represented (independent of the distribution F) as the integral of $(1-e^{-x})^m$ with respect to a gamma density:

$$P\{W_r>m\}=\int_0^{\infty} P\{W_r>m \mid X_{N_r}=x\}\,dP\{X_{N_r}=x\}$$

$$=\int_0^{\infty}(1-e^{-x})^m x^{r-1}e^{-x}\,dx/(r-1)!$$

or

$$P\{W_r=m\}=P\{W_r>m-1\}-P\{W_r>m\}$$

$$=\int_0^{\infty}[(1-e^{-x})^{m-1}-(1-e^{-x})^m]x^{r-1}e^{-x}\,dx/(r-1)!.$$

Notice that the geometric series

$$\sum_{m=1}^{\infty} m[(1-e^{-x})^{m-1}-(1-e^{-x})^m]=\sum_{m=1}^{\infty}(1-e^{-x})^{m-1}$$

$$=\frac{1}{1-(1-e^{-x})}=e^{x};$$

so (as already showed by a different argument) the expectation

$$E(W_r)=\sum_{m=1}^{\infty} mP\{W_r=m\}=\int_0^{\infty}\frac{x^{r-1}\,dx}{(r-1)!}=\infty.$$

(I am trying to avoid integral calculus in this paper, but the preceding argument is too ingenious to omit.)

For large values of r, the gamma can be approximated by a normal density. Neuts [**27**, **28**] used such approximation in the integral expression for $P\{W_r>m\}$ to prove a "law of large numbers" for waiting times: as $r\to\infty$,

$$\frac{\ln(W_r)}{r}\to 1$$

in probability (also: with probability one [20]). By the same method, Neuts gave a "central limit theorem": as $r \to \infty$, the distribution of $(\ln(W_r) - r)/\sqrt{r}$ is asymptotically normal with mean = 0 and variance = 1. There is also a "law of the iterated logarithm" for waiting times [42]. Shorrock ([36] page 222) and Resnick ([33] page 867) include these limit results in more general theorems for inter-record waiting times.

The remarkable similarity between limit theorems for record times and for inter-record waiting times suggests that the wait W_r for the *next* record value is somehow comparable in scale to the sum of all past waiting times $W_1 + W_2 + \cdots + W_{r-1} = N_r - 1$. Resnick [33] connected the limit behaviour of W_r and N_r in the following theorem: with probability one,

$$\limsup_{r\to\infty} \frac{|\ln(W_r) - \ln(N_r)|}{\ln(r)} = 1.$$

Since $W_r \approx e^r$ in some probabilistic limit sense, successive waits must have increasing *median* values, although they all have mode = 1 and mean = ∞. Numeric computations [19] show that

$$\frac{\text{median}(W_{r+1})}{\text{median}(W_r)} \approx e = 2.718\ldots$$

even for $r = 4, 5, 6, 7, 8$. Note that "small r" does not mean "small sample size": Table C shows that fewer than 8 record highs are expected in a sample of size $n = 1000$.

r	2	3	4	5	6	7	8
median(W_r)	4	10	26	69	183	490	1316
med(W_r)/med (W_{r-1})		2.50	2.60	2.65	2.65	2.68	2.69

The end of the preceding section found independent uniform limit distributions for N_r/N_{r+1} and N_{r+1}/N_{r+2}; and hence, as $r \to \infty$, the waiting time ratio

$$\frac{W_{r+1}}{W_r} = \frac{N_{r+2} - N_{r+1}}{N_{r+1} - N_r} = \frac{(N_{r+2}/N_{r+1}) - 1}{1 - (N_r/N_{r+1})}$$

asymptotically has the same distribution as

$$\frac{(1/V) - 1}{1 - U} \quad \text{or} \quad \frac{(1/V) - 1}{U},$$

where U and V are independent random variables uniform over the interval (0, 1). For any $x > 0$,

$$P\left\{\frac{(1/V) - 1}{U} > x\right\} = P\{V < (xU + 1)^{-1}\} = \int_{u=0}^{1} \int_{v=0}^{(xu+1)^{-1}} dv\,du$$

$$= \frac{1}{x}\int_0^1 \frac{x\,du}{xu + 1} = \frac{\ln(1 + x)}{x}.$$

Thus Shorrock [37] obtained the following limit result for ratios of waiting times: for any $x > 0$,

$$P\left\{\frac{W_{r+1}}{W_r} > x\right\} \to \frac{\ln(1 + x)}{x} \quad \text{as} \quad r \to \infty.$$

11. The record value sequence. Little has been said so far about actual record values, i.e., magnitudes. Rather I have focused on frequencies and serial positions of record values.

Results concerning the record frequency R_n in a random sample $X_1, X_2, \ldots, X_n$ or concerning the index number N_r of the r^{th} record value trial are *distribution-free* results: they presume identically distributed and independent (or exchangeable) random variables sampled from a continuous probability distribution; but the distribution function F never appears in the theorems.

The actual record values $X_1 < X_{N_2} < X_{N_3} < X_{N_4} < \cdots$ present a quite contrary situation. Results concerning this record value sequence *per se* definitely depend on the particular distribution F, but do not involve positions in the original random sequence.

Many properties of the record value sequence from continuous distribution F can be stated most conveniently in terms of the transform

$$G(x) = -\ln[1 - F(x)].$$

The derivative $G'(x) = F'(x)/[1 - F(x)]$ is the *failure rate* or *hazard rate* which plays a central role in mathematical theory of reliability (see **[2]** p. 53). Both F and G are strictly increasing functions over their support intervals, so there exist respective inverses F^{-1} and G^{-1},

$$G^{-1}(x) = F^{-1}(1 - e^{-x}).$$

Specific examples will illustrate the distinction between distribution-free results for record value times in random sampling and distribution-dependent results for the record value sequence itself. The theorem of Dwass at the end of Section 8 asserts that, for $0 < \boldsymbol{a} < \boldsymbol{b}$, the count of record value trials indexed between $\boldsymbol{an}$ and $\boldsymbol{bn}$ has asymptotically a Poisson distribution with mean $\ln(\boldsymbol{b}/\boldsymbol{a})$. This result clearly is asymptotic (valid as $n \to \infty$) since there can be at most $(\boldsymbol{b} - \boldsymbol{a})n$ record indices between $\boldsymbol{an}$ and $\boldsymbol{bn}$, while Poisson distribution permits an arbitrarily large frequency. By contrast, for any $\alpha < \beta$ such that $0 < F(\alpha) < F(\beta) < 1$, arbitrarily many of the records $X_1 < X_{N_2} < X_{N_3} < \cdots$ can take values between α and β. Dwass **[10]** proved that this count is *exactly* Poisson distributed, but with mean $G(\beta) - G(\alpha) = -\ln\{[1 - F(\beta)]/[1 - F(\alpha)]\}$, depending on the distribution function F.

The fact that exact distributions are obtained more easily than most limit results for record values is a further contrast to results for record value *times*.

The distribution of the r^{th} record value was required in the preceding section's calculus argument concerning waiting time W_r. Let positive integers $w_1, w_2, \ldots, w_{r-1}$ denote fixed waits, and specify fixed trial numbers $n_2 = 1 + w_1$, $n_3 = n_2 + w_2, \ldots,$ $n_r = n_{r-1} + w_{r-1}$. Consider a random sample $X_1, \ldots,$ $X_{n_2}, \ldots, X_{n_3}, \ldots$ such that

$$\begin{aligned}
&X_1 = x_1 > \text{next } (w_1 - 1) \text{ observations} \\
&X_{n_2} = x_2 > \text{next } (w_2 - 1) \text{ observations} \\
&\vdots \\
&X_{n_r} = x_r.
\end{aligned}$$

Since $X_1, X_2, X_3, \ldots$ are independent and identically distributed with $P\{X_i < x\} = F(x)$, the event above has probability

$$\begin{aligned}
&dF(x_1)[F(x_1)]^{(w_1-1)} \\
\times\ &dF(x_2)[F(x_2)]^{(w_2-1)} \\
&\vdots \\
\times\ &dF(x_{r-1})[F(x_{r-1})]^{(w_{r-1}-1)} \\
\times\ &dF(x_r),
\end{aligned}$$

where $dF(x) = F'(x)dx$ and the derivative F' is a probability density function. For increasing values $x_1 < x_2 < \cdots < x_r$, the event above is equivalent to

$$X_1 = x_1,\ W_1 = w_1,$$
$$X_{N_2} = x_2,\ W_2 = w_2,$$
$$\vdots$$
$$X_{n_{r-1}} = x_{r-1},\ W_{r-1} = w_{r-1},$$
$$X_{N_r} = x_r.$$

Hence summation of the foregoing probability element over all possible waiting times gives the joint probability that $X_1 = x_1$, $X_{N_2} = x_2, \ldots, X_{N_r} = x_r$. This sum can be expressed as a product of geometric series:

$$\sum_{w_1=1}^{\infty} \sum_{w_2=1}^{\infty} \cdots \sum_{w_{r-1}=1}^{\infty} dF(x_1)[F(x_1)]^{(w_1-1)} \cdots dF(x_{r-1})[F(x_{r-1})]^{(w_{r-1}-1)} dF(x_r)$$
$$= dF(x_1) \sum_{w_1=0}^{\infty} [F(x_1)]^{w_1} \cdots dF(x_{r-1}) \sum_{w_{r-1}=0}^{\infty} [F(x_{r-1})]^{w_{r-1}} dF(x_r)$$
$$= \frac{dF(x_1)}{1-F(x_1)} \cdots \frac{dF(x_{r-1})}{1-F(x_{r-1})} dF(x_r)$$
$$= dG(x_1) \cdots dG(x_{r-1})\, dF(x_r).$$

Iterated integration with respect to $x_1, x_2, \ldots, x_{r-1}$ (over the region $x_1 < x_2 < \cdots < x_r$) shows that X_{N_r} has probability element

$$dP\{X_{N_r} = x_r\} = \frac{[G(x_r)]^{r-1}}{(r-1)!}\, dF(x_r).$$

This argument was suggested by Karlin in a textbook exercise ([**21**] pages 267–268; see also [**31**] page 69).

If observations $X_1, X_2, X_3, \ldots$ have *exponential* distribution $F(x) = 1 - e^{-x}$ and $G(x) = -\ln[1 - F(x)] = x$, then the r^{th} record value has *gamma* probability density $x^{r-1}e^{-x}/(r-1)!$, as asserted in the preceding section.

Moreover, for X from any continuous distribution F, it is easy to show that the transformed variable $F(X)$ has uniform distribution on the unit interval ([**21**] page 237) and $G(X)$ has standard exponential distribution. Since G is strictly increasing over its support, the record value trials in random sampling $X_1, X_2, X_3, \ldots$ from F correspond to record value trials in a sample $G(X_1)$, $G(X_2)$, $G(X_3), \ldots$ from the exponential distribution. Thus transformed record value $G(X_{N_r})$ has gamma distribution with r degrees of freedom. For large r the gamma distribution is approximately normal. Thus Resnick [**31**] obtained a "central limit theorem": as $r \to \infty$, the distribution of $(G(X_{N_r}) - r)/\sqrt{r}$ is asymptotically normal with mean $= 0$ and variance $= 1$. The corresponding "strong law of large numbers" asserts that

$$G(X_{N_r})/r \to 1 \quad \textit{with probability one as } r \to \infty.$$

Resnick [**31**] (also see Shorrock [**37**]) gave explicit conditions on the function G which are necessary and sufficient for convergences in probability

$$X_{N_r} - G^{-1}(r) \to 0, \qquad X_{N_r}/G^{-1}(r) \to 1.$$

But also there are conditions under which the last ratio converges to a non-degenerate random variable rather than to a constant. Indeed, Resnick [**31**] characterized the types of limit distributions for record values X_{N_r}; depending on the sampled distribution function F, a record value sequence satisfies exactly one of the following convergences in distribution as $r \to \infty$:

(i)
$$P\left\{\frac{X_{N_r}-G^{-1}(r)}{G^{-1}(r+\sqrt{r})-G^{-1}(r)}<x\right\}\to\mathcal{N}(x)$$

(ii)
$$P\left\{\frac{X_{N_r}}{G^{-1}(r)}<x\right\}\to\begin{cases}0, & x<0\\ \mathcal{N}[\alpha\ln(x)], & x\geqq 0\end{cases}$$

(iii)
$$P\left\{\frac{X_{N_r}-\bar{x}}{\bar{x}-G^{-1}(r)}<x\right\}\to\begin{cases}\mathcal{N}[-\alpha\ln(-x)], & x<0\\ 1, & x\geqq 0\end{cases}$$

where $\mathcal{N}$ denotes the standard normal distribution function, α is a positive constant depending on F, and $\bar{x}$ in (iii) is the upper (necessarily finite) endpoint of the support interval of distribution F.

I omit more precise statements of limit theorems for record value sequences because these results involve complicated conditions on F or G. For details see Resnick **[31]** and also **[9, 16, 32, 34, 35, 36, 37, 38, 39, 40, 43, 44]**.

12. Extreme values and extremal processes. Every random sample $X_1, X_2, \ldots, X_n$ has a *sample maximum* or *upper extreme value* $M_n = \max(X_1, X_2, \ldots, X_n)$. Clearly $M_1 = X_1$ and subsequently every *new* maximum is a record value, i.e., record breaking corresponds to a jump or strict inequality $M_n < M_{n+1}$ in the sequence of sample maxima $X_1 \leqq M_2 \leqq M_3 \leqq M_4 \leqq \cdots$.

Thus a record value sequence can be extracted from a maximal sequence which is already removed from the random sequence; but the converse path is impossible. So maximal sequences might logically be studied before record value sequences and such was the historical precedence, contrary to my arrangement of topics. Gumbel **[18]** has given the early history and extensive bibliography on statistics of extremes, including distinctive applications: for example, using a river's past flood levels (annual maxima of daily observations) to plan dams, etc., with sensible allowance for worse floods in the future (see **[18]**, page 236, for analysis of Mississippi River flooding at Vicksburg).

Historically, studies of sample maxima also have been concomitant to the more general subject of *order statistics* **[8]**, viz., randomly sampled data filed or ranked from smallest value to largest value. In the terms of statistical decision theory, ranked data constitute "sufficient statistics" both for "classical" procedures and for most "nonparametric" methods (of which the tolerance limit construction in Section 6 is one example). Unlike the frequency of record breaking, common sample statistics, such as mean or median annual rainfall, can be computed from ordered data as well as from values in their random sampling sequence (in fact, ranking is generally the fastest way to find the median or other sample percentiles).

In 1943 B. V. Gnedenko **[17]** showed that there are precisely three types of limit distributions for sequences of sample maxima. Gnedenko's three types of *extreme value distributions* correspond exactly to the three limit laws for record values published by Resnick **[31]** in 1973 and mentioned briefly in the preceding section. The limit laws for maxima and for record values from continuous distribution F are linked by Resnick in a duality theorem. That is, the different extreme value distributions partition the space of all continuous distribution functions into disjoint "domains of attraction," and record value distributions determine exactly the same partition. See **[32]** for further comparison of record values and maxima.

* * *

A sequence of sample maxima M_n plotted as a function of sample size n can be regarded (see Shorrock **[39]**) as a discrete-time Markov process with jumps at record value times N_r. Common techniques in stochastic process theory can be used to construct an analogous continuous-time

Markov jump process $M(t)$ such that, for any times $0 \leqq t_1 < t_2 < \cdots < t_n$, the variables $M(t_1) \leqq M(t_2) \leqq \cdots \leqq M(t_n)$ have the same joint distribution as sample maxima $M_1 \leqq M_2 \leqq \cdots \leqq M_n$. The study of such continuous-time *extremal processes* seems to have been originated by Dwass [**10, 11**] and Lamperti [**23**] in the mid-1960s.

Dwass used record value theorems to "motivate some of the results" for extremal processes. But, conversely, Shorrock [**38, 39**] and Resnick [**33, 34, 35**] used continuous-time processes to obtain results for record values. In particular, Resnick [**33**] showed that a continuous-time extremal process $M(t)$ with jump times according to a non-homogeneous Poisson process (with intensity t^{-1}) has a "discrete skeleton" $M(n)$, $n = 1, 2, 3, \ldots$, whose jump times behave precisely as record value times (for n sufficiently large).

Thus Resnick [**33**] used the framework of extremal processes "to give a unified explanation of known limit laws" for record frequencies R_n, for record value trial numbers N_r, and for inter-record waiting times W_r.

This study was supported by the National Research Council of Canada, Grant A8044.

References

Asterisks indicate references most relevant to the theory of record values.

1. American Society for Testing and Materials (1970) "Tentative method for evaluating allowable properties for grades of structural lumber". 1916 Race Street, Philadelphia, Pa. 19103.

2. R. E. Barlow and F. Proschan, Statistical Theory of Reliability and Life Testing, Holt, Rinehart & Winston, New York, 1975.

3.* D. E. Barton and C. L. Mallows, The randomization bases of the problems of the amalgamation of weighted means, J. Roy. Statist. Soc., Ser. B, 23 (1961) 423–433.

4.* ——— and ———, Some aspects of the random sequence, Ann. Math. Statist., 36 (1965) 236–260.

5. British Columbia Department of Agriculture, Climate of British Columbia: tables of temperature, precipitation, and sunshine, Annual reports for 1960–1974. Victoria, B.C., Canada.

6.* K. N. Chandler, The distribution and frequency of record values, J. Roy. Statist. Soc., Ser. B, 14 (1952) 220–228.

7.* F. N. David and D. E. Barton, Combinatorial Chance, Hafner, New York, 1962, pp. 178–183.

8. H. A. David, Order Statistics, Wiley, New York, 1970.

9. L. de Haan and S. I. Resnick, Almost sure limit points of record values, J. Appl. Probability, 10 (1973) 528–542.

10.* M. Dwass, Extremal processes, Ann. Math. Statist., 35 (1964) 1718–1725.

11. ———, Extremal processes, II. Illinois J. Math., 10 (1966) 381–391.

12. W. Feller, An Introduction to Probability Theory and Its Applications, Volume I, 3rd ed. Wiley, New York, 1968.

13.* ———, An Introduction to Probability Theory and Its Applications, Volume II, 2nd ed., Wiley, New York, 1971.

14.* F. G. Foster and A. Stuart, Distribution-free tests in time-series based on the breaking of records, with discussion, J. Roy. Statist. Soc., Ser. B, 16 (1954) 1–22.

15.* F. G. Foster and D. Teichroew, A sampling experiment on the powers of the records tests for trend in a time series, J. Roy. Statist. Soc., Ser. B, 17 (1955) 115–121.

16.* W. Freudenberg and D. Szynal, Limit laws for a random number of record values, Bull. Acad. Polon. Sci. Sér. Sci. Math., Astr. Phys., 24 (1976) 193–199.

17. B. V. Gnedenko, Sur la distribution limite du terme maximum d'une série aléatoire, Ann. of Math., 44 (1943) 423–453.

18. E. J. Gumbel, Statistics of Extremes, Columbia University Press, New York, 1958.

19.* D. Haghighi-Talab and C. Wright, On the distribution of records in a finite sequence of observations, with an application to a road traffic problem, J. Appl. Probability, 10 (1973) 556–571.

20.* P. T. Holmes and W. E. Strawderman, A note on the waiting times between record observations, J. Appl. Probability, 6 (1969) 711–714.

21.* S. Karlin, A First Course in Stochastic Processes, Chapter 9. Academic Press, New York, 1966.

22. W. Kruskal, Statistics, Molière and Henry Adams, American Scientist, 55 (1967) 416–428.

23. J. Lamperti, On extreme order statistics, Ann. Math. Statist., 35 (1964) 1726–1737.

24. M. Loève, Probability Theory, 3rd ed. Van Nostrand, Princeton, N.J., 1963.

25. Meteorological Branch, Annual meteorological summary with comparative data: Vancouver Airport 1937–1964; Vancouver City 1900–1964. Department of Transport, Ottawa, Canada, 1964.

26. R. B. Murphy, Non-parametric tolerance limits, Ann. Math. Statist., 19 (1948) 581–589.

27.* M. Neuts, Waiting times between record observations, J. Appl. Probability, 4 (1967) 206–208.

28.* ———, Probability, Allyn & Bacon, Boston, 1973, pp. 299–300, 432.

29.* G. F. Newell, A theory of platoon formation in tunnel traffic, Operations Res., 7 (1959) 589–598.

30.* A. Rènyi, Theorie des éléments saillants d'une suite d'observations, with summary in English. Colloquium on Combinatorial Methods in Probability Theory, (1962) 104–117. Matematisk Institut, Aarhus Universitet, Denmark.

31.* S. I. Resnick, Limit laws for record values, Stochastic Processes and Their Applications, 1 (1973) 67–82.

32.* ———, Record values and maxima, Annals of Prob., 1 (1973) 650–662.

33.* ———, Extremal processes and record value times, J. Appl. Probability, 10 (1973) 864–868.

34.* ———, Weak convergence to extremal processes, Annals of Prob., 3 (1975) 951–960.

35.* S. I. Resnick and M. Rubinovitch, The structure of extremal processes, Advances in Appl. Probability, 5 (1973) 287–307.

36.* R. W. Shorrock, A limit theorem for inter-record times, J. Appl. Probability, 9 (1972) 219–223. Correction on page 877.

37.* ———, On record values and record times, J. Appl. Probability, 9 (1972) 316–326.

38.* ———, Record values and inter-record times, J. Appl. Probability, 10 (1973) 543–555.

39.* ———, On discrete time extremal processes, Advances in Appl. Probability, 6 (1974) 580–592.

40.* M. M. Siddiqui and R. W. Biondini, The joint distribution of record values and inter-record times, Annals of Prob., 3 (1975) 1012–1013.

41. P. N. Somerville, Tables for obtaining non-parametric tolerance limits, Ann. Math. Statist., 29 (1958) 599–601.

42.* W. E. Strawderman and P. T. Holmes, On the law of the iterated logarithm for inter-record times, J. Appl. Probability, 7 (1970) 432–439.

43.* M. N. Tata, On outstanding values in a sequence of random variables, Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete, 12 (1969) 9–20.

44.* W. Vervaat, Limit theorems for records from discrete distributions, Stochastic Processes and Their Applications, 1 (1973) 317–334.

45. S. S. Wilks, Determination of sample sizes for setting tolerance limits, Ann. Math. Statist., 12 (1941) 91–96.

DEPARTMENT OF MATHEMATICS, AND DEPARTMENT OF HEALTH CARE AND EPIDEMIOLOGY, UNIVERSITY OF BRITISH COLUMBIA, VANCOUVER, BR. COLUMBIA V6T 1W5, CANADA.

THE WILLIAM LOWELL PUTNAM MATHEMATICAL COMPETITION

A. P. HILLMAN, G. L. ALEXANDERSON, L. F. KLOSINSKI

The following results of the thirty-seventh William Lowell Putnam Mathematical Competition, held on December 4, 1976, have been determined in accordance with the governing regulations. This annual contest is supported by the William Lowell Putnam Prize Fund for the Promotion of Scholarship left by Mrs. Putnam in memory of her husband and is held under the auspices of the Mathematical Association of America.

The first prize, five hundred dollars, was awarded to the Department of Mathematics of the **California Institute of Technology,** Pasadena, California. The members of its winning team were Christopher L. Henley, Karl W. Heuer, and Albert L. Wells, Jr.; each was awarded a prize of one hundred dollars.

The second prize, four hundred dollars, was awarded to the Department of Mathematics of