Run-Length Encodings — Corrected Probabilities for Golomb-Codes

Tilo Strutz, University of Rostock

July 25, 2002

In 1966, Secret Agent 00111 [was] back at the Casino again, playing a game of chance, while the fate of mankind hangs in the balance. Each game consists of sequence of favorable events (probability p_1), terminated by the first occurrence of an unfavorable event (Probability $p_0 = 1 - p_1$) ...the game is roulette, and the unfavorable event is the occurrence of 0 ...The problem perplexing the [Secret] Service is how to encode the vicissitudes of a wheel ...Finally a junior clerk who has been reading up on Information Theory, suggests encoding the *run length* between successive unfavorable events. In general, the probability of a run length of r is $p_1^r \cdot p_0$, for $n = 0, 1, 2, 3, \ldots$, which is the familiar *geometric distribution*... [Gol66].

Now in 2002, another young research assistant has computed the probabilities of the geometric distribution $p_1^r \cdot p_0$ using a computer program. He found quite many rounding errors (caused assumedly by using limited calculation tools as logarithm tables and sliding rule) and probably one typo in the run-length dictionaries of Solomon W. Golomb. Table 1 shows the corrected values for $p_r = p_1^r \cdot p_0$.

Reference:
[Gol66] Golomb, S.W.: Run-Length Encodings. *IEEE Transactions on Information Theory*, Vol.12, September 1966, 399–401

	$m=1$		$m=2$		$m=3$		$m=4$	
r	p_r	Codeword	p_r	Codeword	p_r	Codeword	p_r	Codeword
0	1/2	0	0.293	00	0.206	00	0.159	000
1	1/4	10	0.207	01	0.164	010	0.134	001
2	1/8	110	0.146	100	0.130	011	0.113	010
3	1/16	1110	0.104	101	0.103	100	0.095	011
4	1/32	11110	0.073	1100	0.082	1010	0.080	1000
5	1/64	111110	0.052	1101	0.065	1011	0.067	1001
6	1/128	1111110	0.037	11100	0.052	1100	0.056	1010
7	1/256	11111110	0.026	11101	0.041	11010	0.047	1011
8	1/512	111111110	0.018	111100	0.032	11011	0.040	11000
9	1/1024	1111111110	0.013	111101	0.026	11100	0.033	11001
10	1/2048	11111111110	0.009	1111100	0.020	111010	0.028	11010
⋮		⋮		⋮		⋮		⋮

Table 1: Run-length dictionaries for small m ($p_r = p_1^r \cdot p_0$)