Probability of Finding Words in License Plates

Math IA 2016

Matthew Wardynski

Northcoast Prep. Academy

Introduction

California passenger vehicles have standard issue license plates that contain one number, three letters, and three more numbers (in that order), and, rarely, the three letters in a license plate spell a three-letter word. When driving, I often times notice cars that have these three-letter words on them and wonder how common they are. In this investigation, I will determine the probability that the license plate of a California passenger vehicle contains a three-letter word (excluding customized or special license plates). Having determined this probability, I will then explore the applications of my result.

Determining the Theoretical Probability

For the purposes of this investigation, consider the following definitions:

L = number of three-letter possibilities suitable for license plates
w = number of three-letter Scrabble words suitable for license plates
word = a three-letter sequence that is valid in the Official Scrabble Players Dictionary
word plate = a non-custom license plate that contains a three-letter word
car = a passenger vehicle with a standard issue California license plate
p = probability that a car has a word plate, approximated to three decimal places

Ignoring any restrictions, the initial number of three-letter possibilities that make a word would be $26 \times 26 \times 26 = 17{,}576$. According to the Official Scrabble Players Dictionary, there are only 1063 three-letter words. This list can be found on the following page. In the California vehicle code for standard issue license plates, three-letter groups cannot end or begin with an "I," "O," or "Q" because an "I" may be mistaken for the number one, and an "O" or "Q" may be mistaken for the number zero. This would decrease L to $23 \times 26 \times 23 = 13{,}754$, and decrease w to 923 because there are 140 three-letter words in the Scrabble dictionary that start or end with "I" "O" or "Q."

Another factor that decreases both L and w is the list of banned license plates in the CA DMV Vehicle Registration Manual section 4.115, which is displayed to the right. This list includes any three-letter words that could be considered offensive or inappropriate (GOD, JEW, SEX, etc.). So, 143 more three-letter combinations are ruled out, which lowers the total of L down to 13,611. Next, 68 is subtracted off of w because only 68 of the 143 banned words are in the Scrabble dictionary. This makes w = 855.

Restricted License Plate Configuration

ABM	ANO	APE	ARS	ASB	ASS	BAD	BAG	BED
BRA	BUN	BUT	BVD	CHP	CIA	COC	COK	CON
COP	CQC	CQK	CQN	CUL	CUM	CUN	CUR	CUZ
DAG	DAM	DDT	DIC	DIE	DIK	DOA	DUD	DUF
DUM	DUN	FAG	FAN	FAT	FBI	FCK	FKU	FOC
FOK	FQC	FQK	FQU	FUC	FUD	FUG	FUK	FUN
FUX	FUY	GAT	GAY	GEE	GOD	GQD	GUT	HAG
HAM	HEL	HEN	HIC	HIK	HIV	HOG	HOR	HQR
JAP	JAZ	JEW	JIG	KIK	KKK	KOC	KOK	KON
KOX	KQC	KQK	KQN	KQX	KYK	LAY	LSD	MEX
NAG	NGR	NIG	NIP	NUN	OVA	PEA	PEE	PEW
PIG	PIS	POT	POW	PST	PUD	PUS	PYS	QVA
RAG	RAT	RAW	RUT	SAC	SAK	SAM	SEX	SHT
SIF	SIN	SLA	SOB	SOT	SQB	SUE	SUK	SUR
SUX	TIT	TUB	UCK	UPP	UPU	URN	URP	USB
USR	VUE	VUK	VUX	WAD	WOP	WQP	YEP	YID

Note: there are 144 words on this list, but only 143 do not start or end with an "I" "O" or "Q"

Official Scrabble Players Dictionary, Three-Letter Words

AAH AAL AAS ABA ABS ABY ACE ACT ADD ADO ADS ADZ AFF AFT AGA AGE AGO AGS AHA AHI AHS
AID AIL AIM AIN AIR AIS AIT AJI ALA ALB ALE ALL ALP ALS ALT AMA AMI AMP AMU ANA AND
ANE ANI ANT ANY APE APO APP APT ARB ARC ARE ARF ARK ARM ARS ART ASH ASK ASP ASS ATE
ATT AUK AVA AVE AVO AWA AWE AWL AWN AXE AYE AYS AZO BAA BAD BAG BAH BAL BAM BAN BAP
BAR BAS BAT BAY BED BEE BEG BEL BEN BES BET BEY BIB BID BIG BIN BIO BIS BIT BIZ BOA
BOB BOD BOG BOO BOP BOS BOT BOW BOX BOY BRA BRO BRR BUB BUD BUG BUM BUN BUR BUS BUT
BUY BYE BYS CAB CAD CAF CAM CAN CAP CAR CAT CAW CAY CEE CEL CEP CHI CIG CIS COB COD
COG COL CON COO COP COR COS COT COW COX COY COZ CRU CRY CUB CUD CUE CUM CUP CUR CUT
CUZ CWM DAB DAD DAG DAH DAK DAL DAM DAN DAP DAS DAW DAY DEB DEE DEF DEL DEN DEP DEV
DEW DEX DEY DIB DID DIE DIF DIG DIM DIN DIP DIS DIT DOC DOE DOG DOH DOL DOM DON DOR
DOS DOT DOW DRY DUB DUD DUE DUG DUH DUI DUM DUN DUO DUP DYE EAR EAT EAU EBB ECO ECU
EDH EDS EEK EEL EEW EFF EFS EFT EGG EGO EKE ELD ELF ELK ELL ELM ELS EME EMO EMS EMU
END ENG ENS EON ERA ERE ERG ERN ERR ERS ESS EST ETA ETH EVE EWE EYE FAB FAD FAG FAH
FAN FAR FAS FAT FAX FAY FED FEE FEH FEM FEN FER FES FET FEU FEW FEY FEZ FIB FID FIE
FIG FIL FIN FIR FIT FIX FIZ FLU FLY FOB FOE FOG FOH FON FOO FOP FOR FOU FOX FOY FRO
FRY FUB FUD FUG FUN FUR GAB GAD GAE GAG GAL GAM GAN GAP GAR GAS GAT GAY GED GEE GEL
GEM GEN GET GEY GHI GIB GID GIE GIF GIG GIN GIP GIS GIT GNU GOA GOB GOD GOO GOR GOS
GOT GOX GRR GUL GUM GUN GUT GUV GUY GYM GYP HAD HAE HAG HAH HAJ HAM HAO HAP HAS HAT
HAW HAY HEH HEM HEN HEP HER HES HET HEW HEX HEY HIC HID HIE HIM HIN HIP HIS HIT HMM
HOB HOD HOE HOG HOM HON HOO HOP HOT HOW HOY HUB HUE HUG HUH HUM HUN HUP HUT HYP ICE
ICH ICK ICY IDS IFF IFS IGG ILK ILL IMP INK INN INS ION IRE IRK ISM ITS IVY JAB JAG
JAM JAR JAW JAY JEE JET JEU JIB JIG JIN JOB JOE JOG JOT JOW JOY JUG JUN JUS JUT KAB
KAE KAF KAS KAT KAY KEA KEF KEG KEN KEP KEX KEY KHI KID KIF KIN KIP KIR KIS KIT KOA
KOB KOI KOP KOR KOS KUE KYE LAB LAC LAD LAG LAH LAM LAP LAR LAS LAT LAV LAW LAX LAY
LEA LED LEE LEG LEI LEK LET LEU LEV LEX LEY LIB LID LIE LIN LIP LIS LIT LOB LOG LOO
LOP LOT LOW LOX LUD LUG LUM LUN LUV LUX LYE MAC MAD MAE MAG MAM MAN MAP MAR MAS MAT
MAW MAX MAY MED MEG MEH MEL MEM MEN MET MEW MHO MIB MIC MID MIG MIL MIM MIR MIS MIX
MMM MOA MOB MOC MOD MOG MOI MOL MOM MON MOO MOP MOR MOS MOT MOW MUD MUG MUM MUN MUS
MUT MUX MYC NAB NAE NAG NAH NAM NAN NAP NAV NAW NAY NEB NEE NEG NET NEW NIB NIL NIM
NIP NIT NIX NOB NOD NOG NOH NOM NOO NOR NOS NOT NOW NTH NUB NUG NUN NUS NUT OAF OAK
OAR OAT OBA OBE OBI OCA OCH ODA ODD ODE ODS OES OFF OFT OHM OHO OHS OIK OIL OKA OKE
OLD OLE OMA OMS ONE ONO ONS OOF OOH OOT OPA OPE OPS OPT ORA ORB ORC ORE ORG ORS ORT
OSE OUD OUR OUT OVA OWE OWL OWN OWT OXO OXY PAC PAD PAH PAK PAL PAM PAN PAP PAR PAS
PAT PAW PAX PAY PEA PEC PED PEE PEG PEH PEN PEP PER PES PET PEW PHI PHO PHT PIA PIC
PIE PIG PIN PIP PIS PIT PIU PIX PLY POD POH POI POL POM POO POP POS POT POW POX PRO
PRY PSI PST PUB PUD PUG PUL PUN PUP PUR PUS PUT PYA PYE PYX QAT QIS QUA RAD RAG RAH
RAI RAJ RAM RAN RAP RAS RAT RAW RAX RAY REB REC RED REE REF REG REI REM REP RES RET
REV REX REZ RHO RIA RIB RID RIF RIG RIM RIN RIP ROB ROC ROD ROE ROM ROO ROT ROW RUB
RUE RUG RUM RUN RUT RYA RYE RYU SAB SAC SAD SAE SAG SAL SAN SAP SAT SAU SAW SAX SAY
SEA SEC SEE SEG SEI SEL SEN SER SET SEV SEW SEX SHA SHE SHH SHY SIB SIC SIG SIM SIN
SIP SIR SIS SIT SIX SKA SKI SKY SLY SOB SOC SOD SOH SOL SOM SON SOP SOS SOT SOU SOW
SOX SOY SPA SPY SRI STY SUB SUE SUK SUM SUN SUP SUQ SUS SYN TAB TAD TAE TAG TAJ TAM
TAN TAO TAP TAR TAS TAT TAU TAV TAW TAX TEA TEC TED TEE TEG TEL TEN TES TET TEW THE
THO THY TIC TIE TIL TIN TIP TIS TIT TIX TIZ TOD TOE TOG TOM TON TOO TOP TOR TOT TOW
TOY TRY TSK TUB TUG TUI TUM TUN TUP TUT TUX TWA TWO TYE UDO UGH UKE ULU UMM UMP UMS
UNI UNS UPO UPS URB URD URN URP USE UTA UTE UTS VAC VAN VAR VAS VAT VAU VAV VAW VEE
VEG VET VEX VIA VID VIE VIG VIM VIN VIS VOE VOG VOW VOX VUG VUM WAB WAD WAE WAG WAN
WAP WAR WAS WAT WAW WAX WAY WEB WED WEE WEN WET WHA WHO WHY WIG WIN WIS WIT WIZ WOE
WOK WON WOO WOS WOT WOW WRY WUD WYE WYN XIS YAG YAH YAK YAM YAP YAR YAS YAW YAY YEA
YEH YEN YEP YES YET YEW YIN YIP YOB YOD YOK YOM YON YOU YOW YUK YUM YUP ZAG ZAP ZAS
ZAX ZED ZEE ZEK ZEP ZIG ZIN ZIP ZIT ZOA ZOO ZUZ ZZZ

Source: USA - Official Scrabble Players Dictionary 5 (Merriam-Webster) 2014; see Works Cited

So, the calculations are as follows.

$$L = (23 \times 26 \times 23) - 143 \qquad w = 1063 - 140 - 68$$
$$= 13{,}754 - 143 \qquad = 1063 - 208$$
$$= 13{,}611 \qquad = 855$$

Now that L and w have been found, the ratio $\frac{w}{L}$ is used to find the percent of cars that have word plates.

$$\frac{w}{L} = \frac{855}{13{,}611} \approx 6.28\%$$

So, the probability of a randomly chosen car having a word plate is $p = 0.0628$.

Analysis and Applications

The probability that I found, 6.28%, was less than I originally expected. I often see two, three, or even four word plates in a fairly short amount of time. Perhaps there appear to be many word plates because they are more noticeable. For example, someone is probably more likely to notice a license plate that reads "CAT" than one that reads "NKJ." Another element that should be taken into consideration is exactly what is meant by a "word." In this exploration, I define a word to be a valid Scrabble word and used the Scrabble dictionary to find a list of three-letter words. However, this dictionary may not be perfectly accurate for the purposes of this paper because there are some words in it that most people would *not* recognize, and likewise other words that they may recognize that are not included. For example, one of the first words in the three-letter Scrabble dictionary is "aas," which according to Princeton's WordNet is the plural of "aa" which is "a dry form of lava resembling clinkers." Most people (and Microsoft Word 2010) do not recognize "aas" as a word. On the other hand, there are many people who would consider texting terms like "LOL," "WTF," and other abbreviations or acronyms like "ETC" and "DVD" as words. Also, there are certain three-letter proper names that are not found in the Scrabble dictionary (Kim, Tim, etc.) that people may consider words. In order to make the most accurate dictionary for this paper, I would have to make my own list of three-letter words by starting with a list like the Scrabble dictionary, dispose of any words I didn't recognize, and add in the words that I thought were missing from the list. However, for now the original definition of a word is acceptable and I will move forward with my investigation using the theoretical probability of 6.28%.

Now that I have found this theoretical probability, I can use it to answer other practical probability questions that have to do with license plates and see if the results agree with my own experience. If we treat the probability of word plates occurring as a Bernoulli trial, seeing one car is a trial that can be considered a success if the car has a word plate or a failure if it does not. The classic simple example of a Bernoulli trial is flipping a coin where getting heads is considered a success and tails a failure. The license plate trials can be seen as an extremely lopsided coin toss where there is only a 6.28% chance of success and a 93.72% chance of failure.

I see about 50 cars on one roundtrip to school, so I usually see around 3 word plates per trip because 6.28% of 50 is 3.14. However, what is the probability that I will see no word plates? Seeing 50 cars is like repeating the Bernoulli trial 50 times with 6.28% chance of success for each trial. This yields a binomial random variable (Ross). Using this fact, we can find the probability of having a specific number of successes by using the probability mass function of a binomial random variable S having parameters (n, p), given by

$$\mathrm{P}(S = k) = \binom{n}{k} p^k (1-p)^{n-k}, \qquad k = 0, 1, \dots, n$$

and

$$\binom{n}{k} = \frac{n!}{(n-k)!\,k!}$$

where k is the number of successes, n is the number of trials, and p is the probability of success for one trial (Ross). To make sense of this formula, we can look at the individual parts and what they correspond to.

$$\underbrace{\mathrm{P}(S = k)}_{\substack{\text{the probability of} \\ \text{getting } k \text{ successes}}} = \underbrace{\binom{n}{k}}_{\substack{\text{the binomial coefficient,} \\ \text{the number of ways} \\ \text{to arrange } k \\ \text{successes within } n \text{ trials}}} \underbrace{p^k}_{\substack{k \text{ successes} \\ \text{occurring with} \\ \text{probability } p}} \underbrace{(1-p)^{n-k}}_{\substack{n-k \text{ failures} \\ \text{occurring with} \\ \text{probability } 1-p}}$$

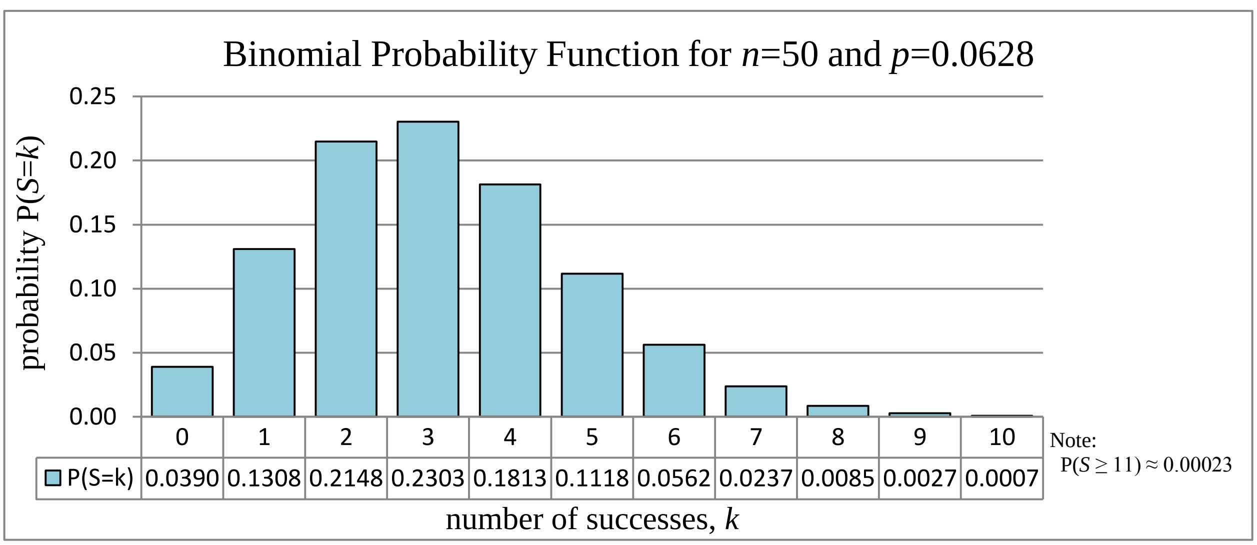
Applying this formula, the probability that we see no word plates, and thus have no successes, is calculated as

$$\mathrm{P}(S = 0) = \binom{50}{0} 0.0628^0 (1 - 0.0628)^{50-0} \approx 0.039049$$

Therefore, there is about a 4% chance that I will not see a word plate on a roundtrip to school.

When looking at the probability function for various values, a graph is helpful. Using the graph and data table below, we can answer other practical questions like, what is the probability that I will see at least 3 word plates out of 50 cars? Since having 3 or more successes is the same as not having 0, 1, or 2 successes, so we can see from the graph that the approximate value is

$$\begin{aligned} \mathrm{P}(S \geq 3) &= 1 - \mathrm{P}(S = 0) - \mathrm{P}(S = 1) - \mathrm{P}(S = 2) \\ &\approx 1 - 0.0390 - 0.1308 - 0.2148 = 0.6154 \end{aligned}$$



k	0	1	2	3	4	5	6	7	8	9	10
P(S=k)	0.0390	0.1308	0.2148	0.2303	0.1813	0.1118	0.0562	0.0237	0.0085	0.0027	0.0007

This probability of 0.6154 leads to the conclusion that 6 out of 10 days (or 3 out of 5 days) I will see 3 or more word plates on a roundtrip to school. This seems reasonable because it agrees with my earlier observation that I often see several word plates in what seems to be a short amount of time. Rather than just relying on this inexact observation, I decided to record empirical observations firsthand. Instead of recording data from cars on a trip to school, each day for a week I went to various locations to collect data from parked cars with standard issue license plates. I recorded the license plate of every car I saw until I reached 50. The following table shows the 50 trials per day with the word plates marked in red. The second row of the table gives the number of word plates per day.

Day	**Day 1**	**Day 2**	**Day 3**	**Day 4**	**Day 5**	**Day 6**	**Day 7**
Success	$S = 3$	$S = 2$	$S = 1$	$S = 6$	$S = 3$	$S = 2$	$S = 4$
1	6JEK858	4JOC412	5WHX026	7EXK834	5LEH891	4PXY801	5MYT778
2	7CLN793	5PRF030	5NBT475	4YNC425	6WDP551	5EEP927	3RKB360
3	3XMK442	6WCY224	6MIA557	6NBE945	6HGA468	6ZZF867	6ZSH087
4	4XTR711	5RJE692	4RRW183	6PSK934	4LGY052	5RAJ670	6WQT168
5	2MOA374	5XCS770	7MTY990	6HRZ296	6HPA551	4TMV252	3GPT224
6	6UUX592	7NGD788	4UXX915	5LRX213	6FBF449	5FGH683	7EEF251
7	7HNG298	5CRY326	5BIJ329	5FVN890	6UJG793	5AWX460	7HVF809
8	6PIN199	4LDZ844	4YXM570	5CTN117	5XHT621	4MCP806	5PJU414
9	6HPV071	7RMG998	6VNA337	4RRA127	5JTC709	4NZG686	5RNF109
10	6STS489	2RHA050	7HFU830	5NFE632	4EKY804	3PQT631	7ASM893
11	5GTN256	4SWM757	7BLM698	4SUH780	5XUC265	5YIJ258	5UVP447
12	5NFK158	4XMG816	7FMA268	5YCY482	7ABA295	5EQF904	6ZOS536
13	5ZMV118	7LCX320	7GJS523	6JYK562	7MEP122	6XXL367	5YPR017
14	6YXM264	6YHE755	7RPF221	6VBU315	6JTP338	6KTN628	6PZS095
15	7PHY127	5ZHL326	4ZEW131	6YIR946	5RPK510	6ZEL057	5CQP421
16	5RPT383	7DWU038	6LUY832	7MJX054	5TBT936	6WSH175	4NIM778
17	7LPW658	4HLL478	5VVF200	7DCT329	6CPW878	5JNA056	5ZPL331
18	5YIG217	5DRD702	6LUY840	4ARB658	6RQP551	4VUT270	4KSM186
19	3TSJ920	7ACM422	6VNA337	6FSE990	6MSD628	5AML676	4KZM034
20	4ZNJ437	6ECT622	4KIU733	5FIX802	7AAL459	5NJP690	6LAP561
21	5TQH254	7AKM010	4BPU518	6CRU106	5KFZ944	7EZE423	6LDB761
22	7JRG296	5WUF270	4WKU006	3LTY570	5YZL126	4MPA029	4BOB446
23	5EZG296	5TMC811	6TWF389	6FHZ827	7EXS892	6CNR971	6REW933
24	5EZG144	5FRS445	6AYG698	7RVL161	4UHP281	5ETD736	6RGR687
25	6LCU497	6XMC897	6UBS336	5MDM481	6RTE231	3TDB576	6GAC740
26	7BQD288	5CQT570	3KIL479	5VQJ262	6CPM168	4WEG240	6WOM629
27	4DSN081	4WVZ591	6SFW189	6WNX013	5NDK854	5SXT347	5ZEB003
28	5RGJ467	5THA112	6LFT087	5XXH958	5HDY090	3WVZ263	6XMR590
29	7NSC663	7DYZ632	4CQD262	7RUP880	4DID051	6LFB276	6ZYA950
30	5TVH166	7NUT693	6RQR687	7KNX520	7CDA750	6SZH579	6TRR309
31	5VWY779	7HLY931	7DWU111	5LVL146	6NSV326	6KTN450	6XMC897
32	6NMH370	4DMV426	4GKE778	5AYE768	5VFG883	6RKC143	6LDF857
33	6VFU096	7BSX582	5SDA164	7JQN180	7BPD181	4LXM344	3VDZ302
34	7HUA976	6ZDY905	6YRH623	6MIZ843	5CHE970	6XCL630	6LVY239
35	7PHY095	7MZK159	6SAV947	7LBJ628	5EFY385	7KTT508	4RSP579
36	3TPG080	6XUL315	6NAN932	7LOB972	5ZDV822	6AQL725	6FNU386
37	6NMG184	6SLV220	5TWM150	4JPX380	5ZRD677	7LZY868	4CRM633
38	6PYJ618	4TEH546	3WOC327	7FWL334	4AME018	6UWM533	6SQG566

39	7KFR227	6SNL244	5LJV784	6EDY452	6UUY499	5UMA924	6RBX890
40	5CIK148	4NLM881	5BZF237	5FQM562	5LSU957	5DTG367	6BFR634
41	6PHM671	7BFU747	7NUZ772	7PAV465	5PUZ269	6JOD397	3EXF572
42	6AOY283	6BLP732	4BJP976	7LHM629	6EXB691	6FYX861	6NCR778
43	6XMU511	7NUE267	4SPK279	5DLD533	7RGN331	7AGP772	5TRV328
44	6DBB263	7DZF675	5VYZ034	7PYW883	6CSU811	6KMY913	7DSV929
45	7KWY842	6FHS554	7CDZ775	6BRJ912	7PYW860	5XPX228	6STM512
46	6BYF918	5BSM199	7LNB144	5YBE599	3JDK591	7RUP912	7ALP697
47	3KDJ199	6LDF857	5JBX788	6GAR842	4GGE099	2SLJ207	6PZS095
48	6DSP568	5BHZ724	3MSF704	6HGA468	7FEB910	4ARB658	5KFB291
49	6ANY206	6UJJ933	7PXU488	6PZA965	4RVR435	3VFK254	5BSM199
50	7MZK063	7AGR378	4LFA350	5MFW868	7LCS870	7BVX257	3MQY001

Over the course of the first 5 days, the table shows 3 days where I saw 3 or more word plates. This agrees with my earlier calculation that $\mathrm{P}(S \geq 3) = 0.6154 \approx \frac{3}{5}$. Overall in my observations, the number of successes, *S*, ranges from 1 to 6 and has a mean of 3.00. This mean is close to the theoretical mean calculated earlier: 3.14. Since my number of samples is not very big, I would expect that the mean would get closer to 3.14 if I increased the number of samples, for example if I recorded 50 trials each day for a month instead of just a week.

In the data in the table, there are some license plates that are observed twice. This may seem to affect the accuracy of the data, but it is practical to see the same car more than once in the course of a week. This investigation is conducted as a probability experiment with replacement; once a car is seen, it is released back into the sample space and can be counted again if chosen. If the trials were done without replacement, each car would disappear from existence once it was seen.

Another interesting question is how probable it is to see two word plates in a row. I have noticed this on the way to school and in the data I collected. For example, on day 4 the word plate 5FIX802 was immediately followed by another word plate, 6CRU106. To find probabilities like this, rather than just looking at the number of successes in *n* trials, we must focus on the time between successes, or the interarrival time. That is, if I just saw a word plate, how many more trials do I need to wait to see another? This number, *T*, corresponds to a geometric random variable. *T* is the number of the trial on which the first success occurs (Mendenhall). Using this fact, we can find the probability of having to wait *n* trials by using the probability mass function of a geometric random variable *T* having parameter *p*, given by

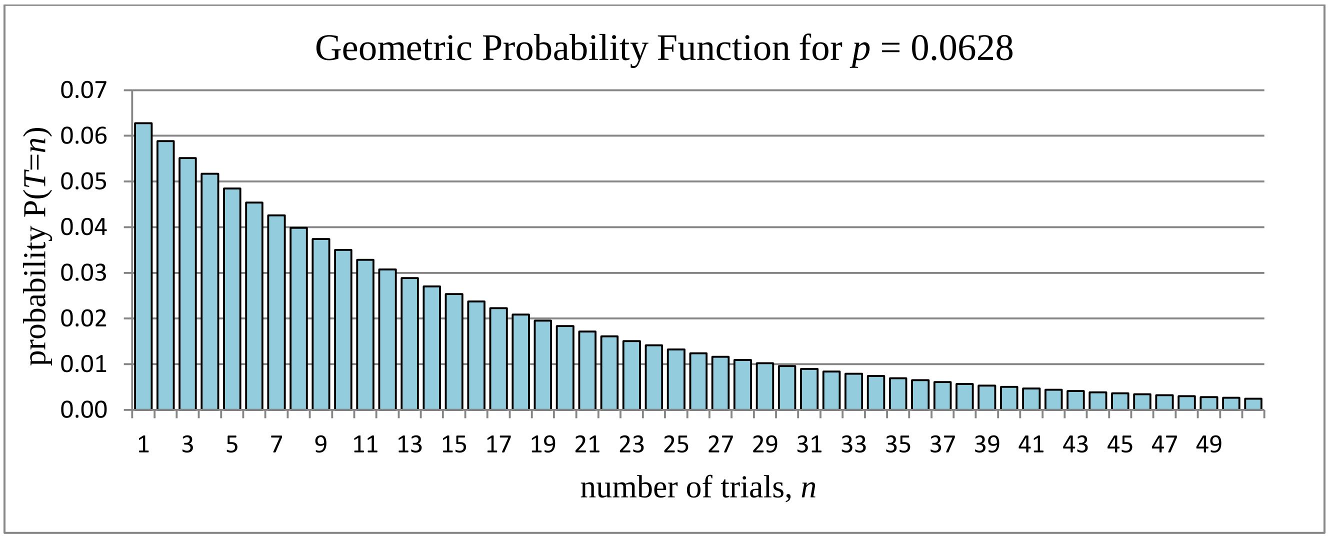
$$\mathrm{P}(T = n) = (1-p)^{n-1}p, \quad n = 1, 2, \ldots$$

where *p* is the probability of success for one trial (Ross, Mendenhall). This makes sense because $1 - p$ is the probability of failure and $n - 1$ is the number of failures, whereas p is the probability of the one and only success. There is also no upper bound on *n* because, theoretically, it is possible that we can keep searching for a word plate indefinitely without success. Also, $n \neq 0$ because *n* is the number of trials away from the first trial we are observing. If $n = 0$ then the interarrival time between the two successes would be 0 and therefore we would be observing the same trial twice.

So, the probability that the interarrival time is 1, or that the two successes come one after the other, is calculated as

$$\mathrm{P}(T = 1) = (1 - 0.0628)^{1-1}0.0628 = 0.0628$$

Therefore, when n = 1 the probability is the same as the original p. At first this seemed confusing to me, but it does make sense that the probability of seeing one word plate after another is the same as the probability of a word plate occurring because we are using the first word plate as a starting place to find the next. In this case, the first word plate is not counted as a trial because it has already been found, so seeing the second license plate is considered the first trial. Thus, it makes sense that when n = 1 the probability is the original p because the probability of a word plate occurring in the first trial is always 0.0628. In order to make further sense of this concept, it is helpful to look at a graph of what the probability for a geometric random variable looks like.



The graph can be used to determine the probabilities in other questions like, what is the probability that the first word plate I see is on the 12th car I pass? From the graph, we see that the probability is roughly 0.03. The probability decreases as the number of trials increase because the process stops once there is one success. That is, it is less likely to go a longer time without getting a success. It may seem counterintuitive that it is more probable to see a word plate on the first trial than on the second, but it is less likely to get to the second trial because 6.28% of the time the success will be found on the first and the process will end.

Since the interarrival time T has a geometric distribution, the mean value, or expected value, is given by $\mu = \frac{1}{p}$ (Mendenhall). Using this equation, we find the expected interarrival time to be 15.9. So, there are about 16 cars in between every word plate. From the 21 successes I observed during my data collection, my interarrival times are 4, 3, 41, 8, 23, 57, 32, 2, 1, 11, 4, 11, 15, 8, 9, 25, 44, 18, 4, 2, and 24. These are the interarrival times if I consider my data as an ongoing Bernoulli process that is not separated by groups of 50. This data set has a median of 11, mean of 16.5, and ranges from 1 to 56. This mean is close to the expected interarrival time of 15.9, which is surprising since my data set is not that large and has such a wide range.

Conclusion

At first when I calculated the theoretical probability *p* of word plates occurring in standard issue California license plates, I did not expect the probability to be so low. However, after applying probability theory, I saw that this theoretical value of *p* lead to results similar to what my own experience with license plates suggested.

Recognizing that finding a word plate on a car is like flipping a lopsided coin helped me model the situation with a Bernoulli trial. Recognizing that seeing 50 cars is like repeating this Bernoulli trial 50 times allowed me to model this situation with a binomial random variable. Using this model, I was able to see how the probabilities were distributed and thus see just how common word plates really are, which was the original intent of this exploration. I was then able to extend my exploration beyond just looking at the number of successes to looking at how the successes themselves were distributed by modeling the interarrival time with a geometric random variable.

It is interesting to see how a simple concept like a Bernoulli trial can lead to so many different results. The probability models used in this paper are helpful for answering questions and conducting experiments about license plate probability, but they can also be used in a variety of applications across different disciplines. There is a wide range of instances where there are two possible outcomes: whether or not a product is defective in manufacturing, whether or not a trait is inherited in genetics, whether or not the patient is cured in medicine, and so on.

In all these applications, the underlying real-life situation is different, but they all can be modeled the same way; mathematical modeling is a powerful tool for understanding the world. Having completed this mathematical exploration, I will continue to observe cars with word plates, but now with a greater understanding of the underlying randomness of the process.

Works Cited

CA Dept. of Motor Vehicles "CA Dept. of Motor Vehicles License Plate Introduction." *CA.gov*. n.d. Web. 18 Mar. 2016.

CA Dept. of Motor Vehicles. "Vehicle Registration Manual, 4.115 Personalized Configurations." *California Code of Regulations, Article 3.* Governmentattic.org, 28 Jan. 2013. Web. 2 Mar. 2016.

Mendenhall, William, et al. *Mathematical Statistics with Applications*. 4th ed. North Scituate, MA: Duxbury, 1990. Print.

Princeton University. "About WordNet." *WordNet*. Princeton University. 2010.

Ross, Sheldon M. *Introduction to Probability Models*. 5th ed. New York: Academic, 1993. Print.

"USA - Official Scrabble Players Dictionary 5 (Merriam-Webster) 2014." *Word Helper.* Richard Loxley Ltd, n.d. Web. 3 Mar. 2016.