read "/Users/bethkupin/Documents/Experimental Math/GouldenJackson/GeneralizedGouldenJackson.txt"

This is the package GeneralizedGouldenJackson, written by Beth Kupin and Debbie Yuster **(1)**

read "/Users/bethkupin/Documents/Experimental Math/GouldenJackson/Frequencies.txt"
read "/Users/bethkupin/Documents/Experimental Math/GouldenJackson/glossary.txt"
$A := [a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z]:$
$M1 := PassageDouble(English, A, SP):$
$M2 := [op(M1), [seq(0, blah = 1..26), 1]]:$

We may assume that every passage begins with a blank space (essentially, no passage begins in the middle of a word), which lets us set the initial letter probabilities to zero for all characters other than SP.

$M := MakeTable(M2, [op(A), SP]):$

Finding the Generating Function for the avoided word [SP,t,h,e]:

$J := map(evalf, ProbDoubleGJ([op(A), SP], \{[SP, t, h, e]\}, x, M1));$

$$-(1. (1.216277405\ 10^{103} x - 1.467619256\ 10^{83} x^{21} + 1.135024668\ 10^{84} x^{20} + 3.929141702\ 10^{100} x^3 + 3.005847673\ 10^{91} x^{14} - 5.009647406\ 10^{99} x^5 - 6.244715911\ 10^{100} x^4 - 4.690517077\ 10^{94} x^{11} - 4.172955980\ 10^{96} x^9 + 5.992161832\ 10^{95} x^{10} - 6.019190194\ 10^{101} x^2 + 1.661354039\ 10^{99} x^6 - 2.082845984\ 10^{77} x^{25} - 3.022001459\ 10^{97} x^7 + 2.813331769\ 10^{96} x^8 - 3.204201707\ 10^{85} x^{19} + 5.075891156\ 10^{93} x^{12} - 2.045993187\ 10^{88} x^{17} - 4.492452262\ 10^{92} x^{13} + 1.049335366\ 10^{79} x^{24} + 9.857119002\ 10^{86} x^{18} + 2.242821867\ 10^{74} x^{26} + 1.345395964\ 10^{74} x^{27} - 2.417648749\ 10^{72} x^{28} + 1.274738199\ 10^{82} x^{22} - 3.478438939\ 10^{90} x^{15} + 4.142898733\ 10^{103} + 3.235479592\ 10^{89} x^{16} + 2.427466828\ 10^{68} x^{29} - 4.604348748\ 10^{80} x^{23})) / (2.926621328\ 10^{103} x - 1.554641177\ 10^{84} x^{21} + 4.315290923\ 10^{85} x^{20} - 6.412104364\ 10^{101} x^3 + 2.361814965\ 10^{92} x^{14} - 3.449955573\ 10^{100} x^5 + 6.926781271\ 10^{100} x^4 - 4.180047658\ 10^{94} x^{11} - 6.628301486\ 10^{96} x^9 + 8.149372949\ 10^{95} x^{10} + 1.276469307\ 10^{103} x^2 - 2.846077441\ 10^{98} x^6 - 1.732312193\ 10^{78} x^{25} - 1.818872311\ 10^{98} x^7 + 2.629858882\ 10^{97} x^8 - 7.555170668\ 10^{86} x^{19} + 1.979716181\ 10^{94} x^{12} - 4.822517580\ 10^{88} x^{17} - 3.346967483\ 10^{93} x^{13} + 5.266042072\ 10^{79} x^{24} + 8.124166002\ 10^{87} x^{18} - 1.773749972\ 10^{74} x^{26} + 5.527515997\ 10^{74} x^{27} - 5.184589721\ 10^{72} x^{28} + 5.670579204\ 10^{82} x^{22} - 1.088670805\ 10^{91} x^{15} + 3.816385343\ 10^{89} x^{16} - 4.142898733\ 10^{103} + 4.942886549\ 10^{68} x^{29} - 1.315916342\ 10^{81} x^{23}) \quad \textbf{(2)}$$

$map(evalf, taylor(ProbDoubleGJ([op(A), SP], \{[SP, t, h, e]\}, x, M), x, 101));$

$$1. + 1.\, x + 1.\, x^2 + 1.\, x^3 + 0.9992162308\, x^4 + 0.9992162308\, x^5 + 0.9991288963\, x^6 \quad \textbf{(3)}$$

$+ 0.9990341113\,x^{7} + 0.9989403663\,x^{8} + 0.9988466147\,x^{9} + 0.9987529643\,x^{10}$
$+ 0.9986592740\,x^{11} + 0.9985656098\,x^{12} + 0.9984719485\,x^{13} + 0.9983782981\,x^{14}$
$+ 0.9982846557\,x^{15} + 0.9981910224\,x^{16} + 0.9980973977\,x^{17} + 0.9980037819\,x^{18}$
$+ 0.9979101748\,x^{19} + 0.9978165765\,x^{20} + 0.9977229870\,x^{21} + 0.9976294063\,x^{22}$
$+ 0.9975358344\,x^{23} + 0.9974422712\,x^{24} + 0.9973487168\,x^{25} + 0.9972551712\,x^{26}$
$+ 0.9971616343\,x^{27} + 0.9970681062\,x^{28} + 0.9969745869\,x^{29} + 0.9968810764\,x^{30}$
$+ 0.9967875746\,x^{31} + 0.9966940816\,x^{32} + 0.9966005974\,x^{33} + 0.9965071220\,x^{34}$
$+ 0.9964136553\,x^{35} + 0.9963201974\,x^{36} + 0.9962267482\,x^{37} + 0.9961333078\,x^{38}$
$+ 0.9960398762\,x^{39} + 0.9959464533\,x^{40} + 0.9958530392\,x^{41} + 0.9957596339\,x^{42}$
$+ 0.9956662373\,x^{43} + 0.9955728495\,x^{44} + 0.9954794704\,x^{45} + 0.9953861001\,x^{46}$
$+ 0.9952927386\,x^{47} + 0.9951993858\,x^{48} + 0.9951060418\,x^{49} + 0.9950127065\,x^{50}$
$+ 0.9949193800\,x^{51} + 0.9948260622\,x^{52} + 0.9947327532\,x^{53} + 0.9946394529\,x^{54}$
$+ 0.9945461614\,x^{55} + 0.9944528787\,x^{56} + 0.9943596047\,x^{57} + 0.9942663394\,x^{58}$
$+ 0.9941730829\,x^{59} + 0.9940798351\,x^{60} + 0.9939865961\,x^{61} + 0.9938933658\,x^{62}$
$+ 0.9938001443\,x^{63} + 0.9937069315\,x^{64} + 0.9936137274\,x^{65} + 0.9935205321\,x^{66}$
$+ 0.9934273456\,x^{67} + 0.9933341678\,x^{68} + 0.9932409987\,x^{69} + 0.9931478383\,x^{70}$
$+ 0.9930546867\,x^{71} + 0.9929615439\,x^{72} + 0.9928684097\,x^{73} + 0.9927752843\,x^{74}$
$+ 0.9926821677\,x^{75} + 0.9925890597\,x^{76} + 0.9924959605\,x^{77} + 0.9924028701\,x^{78}$
$+ 0.9923097883\,x^{79} + 0.9922167153\,x^{80} + 0.9921236511\,x^{81} + 0.9920305955\,x^{82}$
$+ 0.9919375487\,x^{83} + 0.9918445106\,x^{84} + 0.9917514813\,x^{85} + 0.9916584606\,x^{86}$
$+ 0.9915654487\,x^{87} + 0.9914724455\,x^{88} + 0.9913794511\,x^{89} + 0.9912864653\,x^{90}$
$+ 0.9911934883\,x^{91} + 0.9911005200\,x^{92} + 0.9910075604\,x^{93} + 0.9909146096\,x^{94}$
$+ 0.9908216674\,x^{95} + 0.9907287340\,x^{96} + 0.9906358093\,x^{97} + 0.9905428933\,x^{98}$
$+ 0.9904499860\,x^{99} + 0.9903570875\,x^{100} + \mathrm{O}(x^{101})$

evalf(*coeff*(*taylor*(*normal*(*ProbDoubleGJ*([*op*(*A*), *SP*], {[*SP*, *t*, *h*, *e*]}, *x*, *M*)), *x*, 201), *x*, 200));

$$0.9811110978 \quad \textbf{(4)}$$

evalf(*coeff*(*taylor*(*normal*(*ProbDoubleGJ*([*op*(*A*), *SP*], {[*SP*, *t*, *h*, *e*]}, *x*, *M*)), *x*, 301), *x*, 300));

$$0.9719514288 \quad \textbf{(5)}$$

evalf(*coeff*(*taylor*(*normal*(*ProbDoubleGJ*([*op*(*A*), *SP*], {[*SP*, *t*, *h*, *e*]}, *x*, *M*)), *x*, 351), *x*, 350));

$$0.9674037124 \quad \textbf{(6)}$$

This computation takes about 3 minutes on my computer:

t0 := *time*( ) : *evalf*(*coeff*(*taylor*(*normal*(*ProbDoubleGJ*([*op*(*A*), *SP*], {[*SP*, *t*, *h*, *e*]}, *x*, *M*)), *x*, 401), *x*, 400)); *time*( ) − *t0*;

$$4.565$$

$$0.9628772746$$

$$149.719 \quad \textbf{(7)}$$

Another example, with two forbidden words [SP,m,o,m,SP] and [SP,d,a,d,SP].

map(*evalf*, *normal*(*ProbDoubleGJ*([*op*(*A*), *SP*], { [*SP*, *m*, *o*, *m*, *SP*], [*SP*, *d*, *a*, *d*, *SP*]}, *x*, *M*)));

$$\big(3.786448336\ 10^{115}\,x + 8.484938423\ 10^{95}\,x^{21} - 2.115410770\ 10^{97}\,x^{20} + 1.223198918\ 10^{113}\,x^{3} - 1.742513448\ 10^{104}\,x^{14} - 1.195416137\ 10^{112}\,x^{5} - 1.128859738\ 10^{113}\,x^{4} + 2.170274416\ 10^{107}\,x^{11} + 3.248167658\ 10^{108}\,x^{9} - 1.421916357\ 10^{108}\,x^{10} - 1.873861390\ 10^{114}\,x^{2} - 1.594600101\ 10^{111}\,x^{6} + 3.929842444\ 10^{90}\,x^{25} + 6.360451074\ 10^{110}\,x^{7} - 2.646892748\ 10^{109}\,x^{8} + 4.703348493\ 10^{98}\,x^{19} - 2.491847725\ 10^{106}\,x^{12} + 7.845654228\ 10^{100}\,x^{17} + 2.676740829\ 10^{105}\,x^{13} + 1.289744588\ 10^{116} - 1.666249816\ 10^{92}\,x^{24} - 7.039751523\ 10^{99}\,x^{18} - 6.566010696\ 10^{88}\,x^{26} - 1.459285429\ 10^{86}\,x^{27} + 4.479945405\ 10^{85}\,x^{28} - 6.704451525\ 10^{94}\,x^{22} + 6.264951776\ 10^{102}\,x^{15} - 6.130795705\ 10^{101}\,x^{16} - 7.622130596\ 10^{83}\,x^{29} + 7.650751713\ 10^{79}\,x^{30} + 4.584496908\ 10^{93}\,x^{23}\big) \big/ \big(-9.110997547\ 10^{115}\,x - 2.054637732\ 10^{96}\,x^{21} + 6.740049357\ 10^{97}\,x^{20} + 1.996181282\ 10^{114}\,x^{3} - 5.473300132\ 10^{103}\,x^{14} + 1.111657627\ 10^{113}\,x^{5} - 2.352058656\ 10^{113}\,x^{4} - 1.219695441\ 10^{107}\,x^{11} + 1.101719369\ 10^{109}\,x^{9} - 3.791153099\ 10^{107}\,x^{10} - 3.973834475\ 10^{115}\,x^{2} + 3.130106926\ 10^{111}\,x^{6} + 5.533135205\ 10^{90}\,x^{25} + 2.178180100\ 10^{110}\,x^{7} - 7.204717604\ 10^{109}\,x^{8} - 1.809557880\ 10^{99}\,x^{19} - 2.513063381\ 10^{106}\,x^{12} - 6.270351267\ 10^{101}\,x^{17} + 4.972566276\ 10^{105}\,x^{13} + 1.289744588\ 10^{116} - 1.736143659\ 10^{92}\,x^{24} + 3.812549057\ 10^{100}\,x^{18} - 7.664085617\ 10^{88}\,x^{26} - 1.465591381\ 10^{86}\,x^{27} + 4.479945405\ 10^{85}\,x^{28} - 2.156920516\ 10^{94}\,x^{22} - 4.498625413\ 10^{103}\,x^{15} + 7.123153400\ 10^{102}\,x^{16} - 7.622130596\ 10^{83}\,x^{29} + 7.650751713\ 10^{79}\,x^{30} + 3.483568327\ 10^{93}\,x^{23}\big) \quad \textbf{(8)}$$

This result means that after 100 keystrokes the probability that the monkey has typed either "mom" or "dad" is ~10^(-4):

evalf(*coeff*(*taylor*(*normal*(*ProbDoubleGJ*([*op*(*A*), *SP*], { [*SP*, *m*, *o*, *m*, *SP*], [*SP*, *d*, *a*, *d*, *SP*]}, *x*, *M*)), *x*, 101), *x*, 100));

$$0.9990292257 \quad \textbf{(9)}$$

The Hamlet result - what is the probability that the monkey will type "to be or not to be"

map(*evalf*, *normal*(*ProbDoubleGJ*([*op*(*A*), *SP*], { [*SP*, *t*, *o*, *SP*, *b*, *e*, *SP*, *o*, *r*, *SP*, *n*, *o*, *t*, *SP*, *t*, *o*, *SP*, *b*,

$e, SP]\}, x, M)));$

$$\begin{aligned}&\big(1.605388284\ 10^{161}\,x+5.186150814\ 10^{158}\,x^{3}-1.329590857\ 10^{114}\,x^{42}\\&\quad-2.206478435\ 10^{136}\,x^{25}+3.285905953\ 10^{136}\,x^{24}+2.173962706\ 10^{149}\,x^{14}\\&\quad-2.003193662\ 10^{157}\,x^{5}+7.551288246\ 10^{137}\,x^{23}-5.220070794\ 10^{158}\,x^{4}\\&\quad+8.995069477\ 10^{149}\,x^{11}-4.421295194\ 10^{153}\,x^{9}+1.437288300\ 10^{153}\,x^{10}\\&\quad-7.944846609\ 10^{159}\,x^{2}+1.773210204\ 10^{156}\,x^{6}+1.932438522\ 10^{155}\,x^{7}\\&\quad-3.293932007\ 10^{154}\,x^{8}-3.742576166\ 10^{143}\,x^{19}+4.972000731\ 10^{126}\,x^{34}\\&\quad-9.733900871\ 10^{150}\,x^{12}-1.918181482\ 10^{146}\,x^{17}-5.063961599\ 10^{149}\,x^{13}\\&\quad+1.236335455\ 10^{142}\,x^{20}-1.901915712\ 10^{131}\,x^{29}-5.982511344\ 10^{116}\,x^{41}\\&\quad+9.653862835\ 10^{144}\,x^{18}-3.280523529\ 10^{148}\,x^{15}+1.659606443\ 10^{135}\,x^{26}\\&\quad-7.067790835\ 10^{133}\,x^{27}+2.991485316\ 10^{147}\,x^{16}+5.032699959\ 10^{132}\,x^{28}\\&\quad+8.209105219\ 10^{129}\,x^{30}-5.988706580\ 10^{129}\,x^{31}+3.570571306\ 10^{118}\,x^{40}\\&\quad-2.316781430\ 10^{140}\,x^{21}+8.163643883\ 10^{123}\,x^{36}+6.970788117\ 10^{107}\,x^{45}\\&\quad-2.162819787\ 10^{125}\,x^{35}+7.291802438\ 10^{138}\,x^{22}+5.468292891\ 10^{161}\\&\quad-7.724140492\ 10^{127}\,x^{33}+8.923233877\ 10^{128}\,x^{32}-1.503551540\ 10^{120}\,x^{39}\\&\quad+4.173581631\ 10^{121}\,x^{38}-6.251269600\ 10^{122}\,x^{37}-6.944710713\ 10^{111}\,x^{44}\\&\quad+4.081788479\ 10^{113}\,x^{43}\big)\Big/\big(-3.862904607\ 10^{161}\,x+8.463461690\ 10^{159}\,x^{3}\\&\quad-1.329590857\ 10^{114}\,x^{42}-1.582777230\ 10^{136}\,x^{25}-8.926491679\ 10^{137}\,x^{24}\\&\quad+7.237924305\ 10^{149}\,x^{14}+5.019751428\ 10^{158}\,x^{5}-1.156784790\ 10^{138}\,x^{23}\\&\quad-1.040622161\ 10^{159}\,x^{4}-1.436388793\ 10^{153}\,x^{11}+2.851802487\ 10^{154}\,x^{9}\\&\quad+5.858583494\ 10^{153}\,x^{10}-1.684836750\ 10^{161}\,x^{2}+2.180514682\ 10^{157}\,x^{6}\\&\quad-1.579966351\ 10^{156}\,x^{7}-2.261831723\ 10^{155}\,x^{8}-1.002812045\ 10^{145}\,x^{19}\\&\quad-2.700166043\ 10^{126}\,x^{34}-1.063340782\ 10^{151}\,x^{12}-3.183303464\ 10^{147}\,x^{17}\\&\quad+9.227504711\ 10^{150}\,x^{13}+3.867142152\ 10^{143}\,x^{20}+2.025451375\ 10^{133}\,x^{29}\\&\quad-5.982511344\ 10^{116}\,x^{41}+2.014720110\ 10^{146}\,x^{18}-2.502015059\ 10^{149}\,x^{15}\\&\quad+7.680276094\ 10^{135}\,x^{26}+2.451730540\ 10^{134}\,x^{27}+3.579672061\ 10^{148}\,x^{16}\\&\quad-1.304099381\ 10^{134}\,x^{28}-1.856208543\ 10^{132}\,x^{30}+1.133704271\ 10^{131}\,x^{31}\\&\quad+3.570404657\ 10^{118}\,x^{40}-1.266090204\ 10^{142}\,x^{21}+5.254196802\ 10^{123}\,x^{36}\\&\quad+6.970788117\ 10^{107}\,x^{45}-9.610342181\ 10^{124}\,x^{35}+2.206305816\ 10^{140}\,x^{22}\\&\quad+5.468292891\ 10^{161}+1.567836994\ 10^{128}\,x^{33}-5.132797004\ 10^{129}\,x^{32}\\&\quad-1.532570597\ 10^{120}\,x^{39}+4.597287176\ 10^{121}\,x^{38}-6.435979527\ 10^{122}\,x^{37}\\&\quad-6.944710713\ 10^{111}\,x^{44}+4.081788479\ 10^{113}\,x^{43}\big)\end{aligned} \tag{10}$$

After 100 keystrokes, the probability is virtually zero. We know that the probability cannot actually be 0

- the monkey could actually have typed all of Hamlet's famous monologue, but the probability is so small that it does not affect the first 10 digits.

$evalf(coeff(taylor(normal(ProbDoubleGJ([op(A), SP], \{[SP, t, o, SP, b, e, SP, o, r, SP, n, o, t, SP, t, o, SP, b, e, SP]\}, x, M)), x, 101), x, 100));$

$$1.000000000 \tag{11}$$

Even after 300 keystrokes, I have to use a special setting of evalf to get a number that isn't just zero. This computation takes about 3 minutes (on average) on my computer.

$t0 := time() : evalf[30](coeff(taylor(normal(ProbDoubleGJ([op(A), SP], \{[SP, t, o, SP, b, e, SP, o, r, SP, n, o, t, SP, t, o, SP, b, e, SP]\}, x, M)), x, 301), x, 300)); time() - t0;$

$$0.999999999999999999994138275643$$

$$78.344 \tag{12}$$

$evalf[30](1 - \%\%)$

$$5.861724357\ 10^{-21} \tag{13}$$

For comparison, we also have the dumb monkey model. In this model, the monkey types all keys equally at random. This lets us see how much the fact that these are English words affects the probability.

$N1 := \left[seq\left(\left[seq\left(\frac{1}{27}, blah = 1..27\right)\right], blahh = 1..28\right)\right]:$
$N := MakeTable(N1, [op(A), SP]):$

$map(evalf, normal(ProbDoubleGJ([op(A), SP], \{[SP, b, e, t, h, SP]\}, t, N)))$

$$-\frac{27.\left(1.4348907\ 10^{7} + t^{5}\right)}{3.87420489\ 10^{8}\ t + 26.\ t^{6} - 3.87420489\ 10^{8} - 27.\ t^{5}} \tag{14}$$

$t0 := time() : evalf[35](coeff(taylor(normal(ProbDoubleGJ([op(A), SP], \{[SP, t, o, SP, b, e, SP, o, r, SP, n, o, t, SP, t, o, SP, b, e, SP]\}, t, N)), t, 301), t, 300)); time() - t0;$

$$0.99999999999999999999999999337125921$$

$$0.256 \tag{15}$$

$evalf[35](1 - \%\%)$

$$6.62874079\ 10^{-27} \tag{16}$$