A SEATING ARRANGEMENT PROBLEM

Philippe Flajolet
(Version of January 15, 1997)

There are n seats in a row at a luncheonette, and people sit down one at a time at random. They are unfriendly and so never sit next to one another. What is the expected number of persons to sit down? The original problem is due to Freedmann and Shepp, and it appeared as Problem 62-3 in the 1962 volume of *SIAM Review*. There are various alternative formulations. One of them involves fatmen that need more than one stool to sit on. Another one is a simplified description of channel occupation for mobile telephones due to the Math. Center at Bell Labs: there are n consecutive radio channels and stations arrive at random and try to grab a free channel; because of possible interferences, no station occupies a channel next to an already occupied one. What is the expected proportion of occupied channels?

Clearly, the number of occupied seats/channels lie somewhere between $\frac{n}{3}$ and $\frac{n}{2}$. This worksheet explores the way the solution to this and similar problems may be found using the Gfun package. The common schema explored here is: (1) write down an immediate specification of the problem; (2) use the gfun[listtorec] procedure to guess the right differential equation; (3) exploit the results using Maple capabilities for integration and asymptotic expansion.

Basic equations

Let $g(n)$ be the *probability generating function* (PGF) of the number of occupied seats when they are n seats. In the Maple code below, we take implicitly u as the generating variable. If the first individual to arrive occupies seat K, then the number of occupied seats is 1 plus the number of occupied seats in $[1 .. K-2]$ plus the number of occupied seats in $[K+2 .. n]$, as seats $K-1$ and $K+1$ have become unavailable. The subproblems of sizes $K-2$ and $n-K-1$ are of a similar nature. By the randomness assumption, K takes each value in $[1 .. n]$ with equal probability, namely $\frac{1}{n}$. This gives rise to a recurrence on random variables L_n (the number of occupied seats) and on generating functions

```
> L[n]=1+L[K-2]+L[n-K-1],
  Pr(K=k)=1/n;g(n)=u/n*Sum(g(k-2)*g(n-k-1),k=1..n),g(-1)=1,g(0)
  =1,g(1)=u;
```

$$L_n = 1 + L_{K-2} + L_{n-K-1},\ \Pr(K=k) = \frac{1}{n}$$

$$g(n) = \frac{u\left(\sum_{k=1}^{n} g(k-2)\, g(n-k-1)\right)}{n},\ g(-1)=1,\ g(0)=1,\ g(1)=u$$

This recurrence determines the $g(n)$ explicitly and is implemented by the following Maple code:

```
> g:=proc(n) local k; option remember;
```

```
      if n<=0 then 1 elif n=1 then u else
  expand(u/n*convert([seq(g(k-2)*g(n-k-1),k=1..n)],`+`));
      fi
  end:
> seq([j,g(j)],j=0..5);
```

$$[0, 1], [1, u], [2, u], \left[3, \frac{2}{3}u^2 + \frac{1}{3}u\right], [4, u^2], \left[5, \frac{7}{15}u^3 + \frac{8}{15}u^2\right]$$

For instance, when $n = 3$, there is probability $\frac{1}{3}$ that just one seat is occupied: this occurs only if the first person that arrives chooses the middle seat.

We then get the moments by successive differentiation.

```
> subs(u=1,diff([seq(g(j),j=0..20)],u));
```

$$\left[0, 1, 1, \frac{5}{3}, 2, \frac{37}{15}, \frac{26}{9}, \frac{349}{105}, \frac{169}{45}, \frac{11873}{2835}, \frac{7277}{1575}, \frac{157567}{31185}, \frac{233249}{42525}, \frac{11994551}{2027025}, \frac{20791453}{3274425}, \frac{618626159}{91216125}, \frac{307095526}{42567525}, \frac{7545655031}{986792625}, \frac{46426791998}{5746615875}, \frac{1755370057489}{206239658625}, \frac{4368633478802}{488462349375}\right]$$

```
> evalf(subs(u=1,diff([seq(g(j)/j,j=1..40)],u)),5);
```

[1., .50000, .55556, .50000, .49333, .48148, .47483, .46944, .46533, .46203, .45933, .45708, .45518, .45355, .45213, .45089, .44980, .44883, .44796, .44718, .44648, .44583, .44525, .44471, .44421, .44376, .44333, .44294, .44257, .44223, .44191, .44161, .44133, .44107, .44082, .44058, .44036, .44015, .43995, .43976]

This suggests that the mean occupation ration could be, for n large, asymptotic to a constant with approximate value 44%.

The mean occupancy ratio

The easiest is to try a heuristic approach. As the recurrences for moments are linear, it is reasonable to expect them to be of the holonomic type. We thus compute a few dozen initial values and try to guess a recurrence with gfun[listtorec].

```
> with(gfun);
```

[*Laplace*, *algebraicsubs*, *algeqtodiffeq*, *algeqtoseries*, *algfuntoalgeq*, *borel*, *cauchyproduct*, *diffeq**diffeq*, *diffeq+diffeq*, *diffeqtorec*, *guesseqn*, *guessgf*, *hadamardproduct*, *holexprtodiffeq*, *invborel*, *listtoalgeq*, *listtodiffeq*, *listtohypergeom*, *listtolist*, *listtoratpoly*, *listtorec*, *listtoseries*, *listtoseries/Laplace*, *listtoseries/egf*, *listtoseries/lgdegf*, *listtoseries/lgdogf*, *listtoseries/ogf*, *listtoseries/revegf*, *listtoseries/revogf*, *maxdegcoeff*, *maxdegeqn*, *maxordereqn*, *mindegcoeff*, *mindegeqn*, *minordereqn*, *optionsgf*, *poltodiffeq*, *poltorec*, *ratpolytocoeff*, *rec**rec*, *rec+rec*, *rectodiffeq*, *rectoproc*, *seriestoalgeq*, *seriestodiffeq*, *seriestohypergeom*, *seriestolist*, *seriestoratpoly*, *seriestorec*, *seriestoseries*]

```
> rec:=listtorec(subs(u=1,diff([seq(g(j),j=0..25)],u)),u(n));
```

$rec := [\{-2\,u(n) + (-n+1)\,u(n+1) + (2\,n+4)\,u(n+2) + (-n-3)\,u(n+3), u(0) = 0,$

$u(1)=1,\ u(2)=1\},\ ogf]$

The recurrence transforms into a differential equation by means of gfun[rectodiffeq]

```
> ode:=rectodiffeq(op(1,rec),u(n),Y(z));
```

$$ode := \{(2z^2-2z)\,Y(z)+(z^2-2z+1)\left(\frac{\partial}{\partial z}Y(z)\right)-1,\ Y(0)=0\}$$

Now, we are sure of the existence of a closed form for the generating function of averages, since any ODE of order 1 is solvable by quadratures. The dsolve command of Maple does the job:

```
> M1_z:=factor(op(2,dsolve(ode,Y(z))));
```

$$M1\_z := \frac{1}{2}\frac{(e^{(2z)}-1)\,e^{(-2z)}}{(-1+z)^2}$$

We can check consistency with known values

```
> series(M1_z,z=0,30);
```

$$z+z^2+\frac{5}{3}z^3+2z^4+\frac{37}{15}z^5+\frac{26}{9}z^6+\frac{349}{105}z^7+\frac{169}{45}z^8+\frac{11873}{2835}z^9+\frac{7277}{1575}z^{10}+\frac{157567}{31185}z^{11}+$$
$$\frac{233249}{42525}z^{12}+\frac{11994551}{2027025}z^{13}+\frac{20791453}{3274425}z^{14}+\frac{618626159}{91216125}z^{15}+\frac{307095526}{42567525}z^{16}+$$
$$\frac{7545655031}{986792625}z^{17}+\frac{46426791998}{5746615875}z^{18}+\frac{1755370057489}{206239658625}z^{19}+\frac{4368633478802}{488462349375}z^{20}+$$
$$\frac{365468962351379}{38979295480125}z^{21}+\frac{1001316834294748}{102088631019375}z^{22}+\frac{45904893141293831}{4482618980214375}z^{23}+$$
$$\frac{68644124105752304}{6431583754220625}z^{24}+\frac{1368973529334924634 3}{123272021955895312 5}z^{25}+\frac{2218738100762970969 4}{1923043542511966875}z^{26}+$$
$$\frac{703062131058975189 49}{5873549281427953125}z^{27}+\frac{3794345053488099169 16}{3059387453996310937 5}z^{28}+$$
$$\frac{48314159293838846231 2291}{37643577344671751578125}z^{29}+O(z^{30})$$

We can be also quite sure that the process makes sense if we compare as well with values that haven't been used at all in the "guessing" phase:

```
> subs(u=1,diff([seq(g(j),j=26..29)],u));
```

$$\left[\frac{22187381007629709694}{1923043542511966875},\frac{70306213105897518949}{5873549281427953125},\frac{379434505348809916916}{30593874539963109375},\right.$$
$$\left.\frac{4831415929383884623122 91}{37643577344671751578125}\right]$$

The generating function of expectations is meromorphic with only a pole at $z=1$. In order to analyse the coefficients of the explicit solution found, we examine the singular expansion at the double pole $z=1$:

```
> map(normal,series(M1_z,z=1,4));
```

$$\frac{1}{2}(\mathbf{e}^2-1)\,\mathbf{e}^{(-2)}(-1+z)^{-2}+\mathbf{e}^{(-2)}(-1+z)^{-1}-\mathbf{e}^{(-2)}+\frac{2}{3}\mathbf{e}^{(-2)}(-1+z)+\mathrm{O}((-1+z)^2)$$

By the principles of singularity analysis, it is enough to expand the singular part. This shows that the mean number of occupied seats satisfies the approximate formula

```
> m1:=1/2*(1-exp(-2))*(n+1)-exp(-2); evalf(m1,20);
  C1:=evalf(coeff(m1,n,1)):
```

$$m1 := \frac{1}{2}(1-\mathbf{e}^{(-2)})(n+1)-\mathbf{e}^{(-2)}$$

$$.43233235838169365406\,n+.29699707514508096217$$

The asymptotic approximation is in fact extremely good and is about $10^{(-26)}$ for $n=30$

```
> for j from 0 to 30 by 5 do
  j,evalf(subs(u=1,diff(g(j),u))-subs(n=j,m1),30) od;
```

$$0,\ -.296997075145080962159000757542$$

$$5,\ .00800779961311743424266464655 5$$

$$10,\ -.319864455718522868582236 5\ 10^{-5}$$

$$15,\ .1490022835710890168 8\ 10^{-9}$$

$$20,\ -.157792286723990\ 10^{-14}$$

$$25,\ .536622242\ 10^{-20}$$

$$30,\ -.726\ 10^{-26}$$

In particular the constant found empirically to be close to 0.44 is precisely

```
> 1/2*(1-exp(-2))=evalf(1/2*(1-exp(-2)),30);
```

$$\frac{1}{2}-\frac{1}{2}\mathbf{e}^{(-2)}=.432332358381693654053000252514$$

Distributional analysis

Whenever possible in analysis of algorithms, one should try to determine how characteristic the average case is. We show now that the standard deviation of the distribution of the number of occupied seats/channels is $\mathrm{O}(\sqrt{n})$. By the Markov-Chebyshev inequalities, this means that, for large n, almost all configurations must be close to the average predicted value.

First, we have access to the second (factorial) moments by a double differentiation.

```
> l2:=subs(u=1,diff([seq(g(j),j=0..25)],u,u));
```

$$l2 := \left[0, 0, 0, \frac{4}{3}, 2, \frac{58}{15}, \frac{50}{9}, \frac{278}{35}, \frac{158}{15}, \frac{38494}{2835}, \frac{26716}{1575}, \frac{646588}{31185}, \frac{211558}{8505}, \frac{945638}{32175}, \frac{37415492}{1091475},\right.$$
$$\frac{3606958324}{91216125}, \frac{274743086}{6081075}, \frac{555662598058}{10854718875}, \frac{66172924294}{1149323175}, \frac{884549001398}{13749310575},$$
$$\frac{11636213227258}{162820783125}, \frac{21526530167294}{272582485875}, \frac{8866653505362722}{102088631019375}, \frac{4263253050498241 22}{4482618980214375},$$
$$\left.\frac{60652124252469934}{584689432201875}, \frac{661767382419713052 2}{58700962836140625}\right]$$

The function gfun[listtorec] aims at detecting ("guessing") plausible linear recurrences with polynomial coefficients. It uses two parameters gfun[maxordereqn], gfun[maxdegcoeff], with default values 3 and 4 respectively, so that it is tuned for quick discovery of simple recurrences.

```
> gfun['maxordereqn'],gfun['maxdegcoeff'];
  rec:=listtorec(l2,u(n));
```

$$3, 4$$

$$rec := FAIL$$

The control parameters can be set to higher values.

```
> gfun['maxordereqn']:=8; gfun['maxdegcoeff']:=6;
```

$$gfun_{maxordereqn} := 8$$

$$gfun_{maxdegcoeff} := 6$$

We can now determine the right recurrence (we need more values):

```
> l2:=subs(u=1,diff([seq(g(j),j=0..55)],u,u)):
> rec2:=listtorec(l2,u(n));
```

$$\begin{aligned}rec2 := \Big[\{ &(48+24\,n)\,\mathrm{u}(n) + (-32+22\,n+18\,n^2)\,\mathrm{u}(n+1)\\ &+ (28-61\,n-18\,n^2+3\,n^3)\,\mathrm{u}(n+2) + (-12+13\,n-18\,n^2-7\,n^3)\,\mathrm{u}(n+3)\\ &+ (-252-87\,n+14\,n^2+5\,n^3)\,\mathrm{u}(n+4) + (20-41\,n-24\,n^2-3\,n^3)\,\mathrm{u}(n+5)\\ &+ (480+296\,n+60\,n^2+4\,n^3)\,\mathrm{u}(n+6) + (-280-166\,n-32\,n^2-2\,n^3)\,\mathrm{u}(n+7),\\ &\mathrm{u}(5)=\frac{58}{15}, \mathrm{u}(6)=\frac{50}{9}, \mathrm{u}(0)=0, \mathrm{u}(1)=0, \mathrm{u}(2)=0, \mathrm{u}(3)=\frac{4}{3}, \mathrm{u}(4)=2 \}, ogf \Big]\end{aligned}$$

The recurrence is of order 7. The generating function satisfies a differential equation of order 3 with coefficients of degree 7 (!!). It is rather remarkable that the dsolve command of Maple can solve this explicitly.

```
> ode2:=rectodiffeq(op(1,rec2),u(n),Y(z));
```

$$\begin{aligned}ode2 := \{ &\mathrm{Y}(0)=0, \mathrm{D}(Y)(0)=0, (D^{(3)})(Y)(0)=8, (D^{(6)})(Y)(0)=4000,\\ &(D^{(4)})(Y)(0)=48, (D^{(2)})(Y)(0)=0, (D^{(5)})(Y)(0)=464,\\ &(-1440\,z^6+1080\,z^5-1620\,z^4+720\,z^3)\,\mathrm{Y}(z)\\ &+ (-420\,z^5-120\,z^6-720\,z^7+240\,z^2+900\,z^4+120)\left(\frac{\partial}{\partial z}\mathrm{Y}(z)\right)\\ &+ (-540\,z^7+810\,z^6-720\,z^5+930\,z^4-360\,z^3-120\,z)\left(\frac{\partial^2}{\partial z^2}\mathrm{Y}(z)\right)\\ &+ (-90\,z^7+210\,z^6-150\,z^5+90\,z^4-120\,z^3+60\,z^2)\left(\frac{\partial^3}{\partial z^3}\mathrm{Y}(z)\right)\}\end{aligned}$$

Note that a simple rational solution is detected by dsolve. This entails a reduction of order, and a complete algorithm exists for order 2 (in fact that Maple succeeds in bypassing here).

```
> infolevel[dsolve]:=5: M2_z:=factor(op(2,dsolve(ode2,Y(z))));
dsolve/diffeq/polylinearODE:   checking Euler equation
dsolve/diffeq/expsols:   trying exponential solutions
dsolve/diffeq/expsols:   rational solutions partially successful. Result(s)=
```

```
(1-3*z)/(-1+z)^3
dsolve/diffeq/expsols_solvericcati:   all solutions by polynomial part
dsolve/diffeq/expsols:   expon. solutions partially successful. Result(s) =
exp(Int((-2*z^2+z-2)/(z^2-z),z)), exp(Int((-4*z^2-2*z)/(-1+z^2),z))
```

$$M2\_z := \frac{1}{4}\,\frac{1-3\,z+4\,\mathbf{e}^{(-2\,z)}\,z^2-\mathbf{e}^{(-4\,z)}-\mathbf{e}^{(-4\,z)}\,z}{(-1+z)^3}$$

The singular part at $z = 1$ is analysed. For the variance this leads to approximate formulae.

```
> Moment2_sing:=map(simplify,series(M2_z+M1_z,z=1,3));
```

$$Moment2\_sing := \left(-\frac{1}{2}-\frac{1}{2}\,\mathbf{e}^{(-4)}+\mathbf{e}^{(-2)}\right)(-1+z)^{-3}-\frac{1}{4}\,(-7+\mathbf{e}^4+2\,\mathbf{e}^2)\,\mathbf{e}^{(-4)}\,(-1+z)^{-2}-$$
$$3\,\mathbf{e}^{(-4)}\,(-1+z)^{-1}+\mathrm{O}((-1+z)^0)$$

```
> var_n_asympt:=factor(collect(expand(convert([seq((-1)^j*coeff
  (Moment2_sing,z-1,j)*binomial(n-j-1,-j-1),j=-3..-1)],`+`)-m1^
  2),n,simplify));
```

$$var\_n\_asympt := \mathbf{e}^{(-4)}\,(n+3)$$

Again, the approximations are extremely good:

```
> evalf(var_n_asympt,30);
```

$$.0183156388887341802937180212732\,n+.0549469166662025408811540638196$$

```
> evalf(subs(n=50,var_n_asympt),30);
  evalf(subs(u=1,(diff(g(50),u,u)+diff(g(50),u)-diff(g(50),u)^2
  )),30);
```

$$.970728861102911555567055127480$$

$$.970728861102911555567055127482$$

We have in passing obtained a **Theorem**. *In the random seating problem, the variance of the number of occupied seats when there are n seats is asymptotic to*

$$\frac{n+3}{\mathbf{e}^4} = .0183156388887341802937180212732\,n + .054946916666202540881154063820.$$

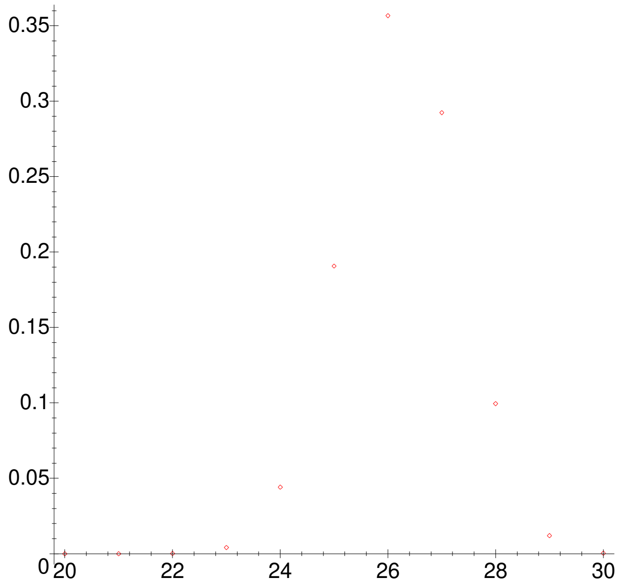
This result seems to be new (!). Convergence is extremely fast so that this formula is highly accurate. The standard deviation is found to be only about $\frac{\sqrt{n}}{7}$.

Distribution. A mean that is $\mathrm{O}(n)$ and a standard deviation that is $\mathrm{O}(\sqrt{n})$ entail that the distribution is concentrated around its mean with high probability. This also suggests that the distributions of the number of occupied seats could be asymptotically Gaussian.

```
> distr:=sort(map(proc(e) [op([2,2],e),op(1,e)]
  end,[op(evalf(g(60),4))]),proc(x,y) evalb(op(1,x)<op(1,y))
  end):
> linalg[transpose](matrix(distr));
```

$$[20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30]$$
$$[.2868\ 10^{-9}, .8641\ 10^{-6}, .0001270, .004097, .04416, .1907, .3567, .2924, .09950, .01200, .0003348]$$

```
> plot(distr,style=POINT);
```



In 65% of the cases, the occupation is either 26 or 27; the probability of an extremely bad assignment (20 seats out of 60) is only about $3\ 10^{(-9)}$. In fact, a Gaussian law can be proved by adapting the bivariate analysis of patterns in binary search trees by Flajolet, Martinez, and Gourdon (1996).

Fatter men

The approach extends to the case where fatmen need b seats on each side. The earlier case was $b = 1$. We consider here $b = 2$. This time, the number of occupied seats lies somewhere between $\frac{n}{3}$ and $\frac{n}{4}$. Now, we let gb be the probability generating function (PGF) with u the generating variable. The following procedure computes gb for a given size n and the parameter b.

```
> gb:=proc(n,b) local k; option remember;
      if n<=0 then 1 elif n<=b then u else
  expand(u/n*convert([seq(gb(k-b-1,b)*gb(n-k-b,b),k=1..n)],`+`)
  )
      fi
  end:
```

The probability generating functions are now:

```
> seq([j,gb(j,2)],j=0..6);
```

$$[0, 1], [1, u], [2, u], [3, u], \left[4, \frac{1}{2}u^2 + \frac{1}{2}u\right], \left[5, \frac{4}{5}u^2 + \frac{1}{5}u\right], [6, u^2]$$

For instance, for $n = 4$, we have 2 occupied seats if the first arrival is on a side (this has probability $\frac{1}{2}$), which leaves the opposite seat available, and 1 occupied seat otherwise.

The moments are obtained by differentiation of the PGF:

```
> subs(u=1,diff([seq(gb(j,2),j=0..25)],u));
```

$$\left[0, 1, 1, 1, \frac{3}{2}, \frac{9}{5}, 2, \frac{16}{7}, \frac{103}{40}, \frac{128}{45}, \frac{1091}{350}, \frac{95}{28}, \frac{55453}{15120}, \frac{322841}{81900}, \frac{371891}{88200}, \frac{363773}{81000}, \frac{7493767}{1572480}, \frac{294729143}{58476600}, \frac{4935980947}{928746000}, \frac{43834851251}{7842744000}, \frac{63303435199}{10795680000}, \frac{325638002203}{53049971520}, \frac{24184493623033}{3771239472000}, \frac{1845637488379687}{275986161360000}, \frac{42103935641621927}{6047696753280000}, \frac{38574335284278439}{5330501946000000}\right]$$

```
> evalf(subs(u=1,diff([seq(gb(j,2)/j,j=1..35)],u)),5);
```

[1., .50000, .33333, .37500, .36000, .33333, .32653, .32188, .31605, .31171, .30844, .30563, .30322, .30118, .29940, .29785, .29648, .29526, .29417, .29319, .29230, .29149, .29076, .29008, .28946, .28889, .28836, .28786, .28740, .28698, .28658, .28620, .28585, .28551, .28520]

This suggests an occupation ratio of about 28%, now.

Like before, we can guess a differential equation and attempt to solve it.

```
> lb1:=subs(u=1,diff([seq(gb(j,2),j=0..35)],u)):
  gfun['maxordereqn']:=5; gfun['maxdegcoeff']:=5;
  recb:=listtorec(lb1,u(n));
```

$$gfun_{maxordereqn} := 5$$

$$gfun_{maxdegcoeff} := 5$$

$$recb := [\{u(0)=0, u(1)=1, u(2)=1, u(3)=1, 2\,u(n) - 2\,u(n+1) + (n+2)\,u(n+2) + (-6-2\,n)\,u(n+3) + (n+4)\,u(n+4)\}, ogf]$$

```
> rectodiffeq(op(1,recb),u(n),Y(z));
```

$$\left\{Y(0)=0, (-4\,z^3+4\,z^2)\,Y(z) + (-2\,z^2+4\,z-2)\left(\frac{\partial}{\partial z}Y(z)\right)+2\right\}$$

The differential equation is of first order, hence again solvable by quadratures

```
> solb:=dsolve(",Y(z));
dsolve/diffeq/dsol1:   -> first order, first degree methods :
dsolve/diffeq/dsol1:   trying linear bernoulli
dsolve/diffeq/linearsol:   solving 1st order linear d.e.
dsolve/diffeq/dsol1:   linear bernoulli successful
```

$$solb := Y(z) = -\frac{1}{2}\,\frac{I\,\mathrm{erf}(I\,z+I)\,\pi\,\mathbf{e}^{(-(1+z)^2)} - I\,\mathbf{e}^{(-z\,(z+2))}\,\pi\,\mathrm{erf}(I)\,\mathbf{e}^{(-1)}}{(z^2-2\,z+1)\sqrt{\pi}}$$

This time the solution involves the Gaussian error function (erf). The solution is singular at $z=1$, with an apparent double pole.

```
> singb:=series(op(2,solb),z=1,3);
```

$$singb := -\frac{1}{2}\,\frac{I\,\mathrm{erf}(2\,I)\,\pi\,\mathbf{e}^{(-4)} - I\,\mathbf{e}^{(-3)}\,\pi\,\mathrm{erf}(I)\,\mathbf{e}^{(-1)}}{\sqrt{\pi}}(-1+z)^{-2} - \frac{1}{2}\,\frac{-4\,I\,\mathrm{erf}(2\,I)\,\pi\,\mathbf{e}^{(-4)} - 2\sqrt{\pi}\,\mathbf{e}^4\,\mathbf{e}^{(-4)} + 4\,I\,\mathbf{e}^{(-3)}\,\pi\,\mathrm{erf}(I)\,\mathbf{e}^{(-1)}}{\sqrt{\pi}}(-1+z)^{-1} - \frac{1}{2}\,\frac{7\,I\,\mathrm{erf}(2\,I)\,\pi\,\mathbf{e}^{(-4)} + 4\sqrt{\pi}\,\mathbf{e}^4\,\mathbf{e}^{(-4)} - 7\,I\,\mathbf{e}^{(-3)}\,\pi\,\mathrm{erf}(I)\,\mathbf{e}^{(-1)}}{\sqrt{\pi}} + O(-1+z)$$

```
> c2:=factor(simplify(coeff(singb,z-1,-2)));
  c1:=factor(simplify(coeff(singb,z-1,-1))); C2:=evalf(c2):
```

$$c2 := \frac{1}{2} I \sqrt{\pi}\, \mathbf{e}^{(-4)} (-\mathrm{erf}(2\,I) + \mathrm{erf}(I))$$

$$c1 := -\frac{I\,(-2\,\mathrm{erf}(2\,I)\,\mathbf{e}^{(-4)}\,\pi + I\sqrt{\pi} + 2\,\mathrm{erf}(I)\,\mathbf{e}^{(-4)}\,\pi)}{\sqrt{\pi}}$$

This corresponds to an asymptotic form for the first moment

```
> mb1:=c2*(n+1)-c1; evalf(mb1,30);
```

$mb1 :=$

$$\frac{1}{2} I \sqrt{\pi}\, \mathbf{e}^{(-4)} (-\mathrm{erf}(2\,I) + \mathrm{erf}(I))\,(n+1) + \frac{I\,(-2\,\mathrm{erf}(2\,I)\,\mathbf{e}^{(-4)}\,\pi + I\sqrt{\pi} + 2\,\mathrm{erf}(I)\,\mathbf{e}^{(-4)}\,\pi)}{\sqrt{\pi}}$$

$$.274550987725779445633557292602\, n + .372754938628897228167786463010$$

Once more, the asymptotic approximation is extremely good, even for relatively small values of n.

```
> for j from 0 to 30 by 5 do
  j,evalf(subs(u=1,diff(gb(j,2),u))-subs(n=j,mb1),30) od;
```

0, -.372754938628897228167786463010

5, .0544901227422055436644270739822

10, -.00112195874383454164621653188 78

15, $.49368424357786868788393222\ 10^{-5}$

20, $.18260946840836293021922\ 10^{-8}$

25, $-.33063392768593603947 8\ 10^{-10}$

30, $.41439988059320442 2\ 10^{-13}$

This last example demonstrates the interest of preserving initial conditions whenever possible. The way Gfun and Maple manage them consistently is especially useful here.

Fatter men, even!

We follow the same schema and consider finally the situation where 3 seats/channels are unavailable next to an occupied seat. The number of occupied seats must now lie between n/4 and n/5.

```
> evalf(subs(u=1,diff([seq(gb(j,3)/j,j=1..35)],u)),5);
```

[1., .50000, .33333, .25000, .28000, .27778, .26531, .25000, .24444, .24133, .23841, .23505, .23223, .22997, .22809, .22641, .22491, .22357, .22238, .22131, .22035, .21946, .21866, .21792, .21725, .21662, .21604, .21550, .21500, .21453, .21410, .21369, .21330, .21294, .21260]

This suggests an occupation ratio of about 21%.

```
> lb1:=subs(u=1,diff([seq(gb(j,3),j=0..35)],u)):
  gfun['maxordereqn']:=6; gfun['maxdegcoeff']:=3;
  recb:=listtorec(lb1,u(n));
```

$$gfun_{maxordereqn} := 6$$

$$gfun_{maxdegcoeff} := 3$$

$$recb := [\{ 2\,u(n) - 2\,u(n+1) + (n+3)\,u(n+3) + (-8-2\,n)\,u(n+4) + (n+5)\,u(n+5),\ u(0)=0,\ u(1)=1,\ u(2)=1,\ u(3)=1,\ u(4)=1 \},\ ogf]$$

```
> gfun['maxdegcoeff']:=1; rectodiffeq(op(1,recb),u(n),Y(z));
```

$$gfun_{maxdegcoeff} := 1$$

$$\{ (-12\,z^4 + 12\,z^3)\,Y(z) + (-6\,z^2 + 12\,z - 6)\left(\frac{\partial}{\partial z} Y(z)\right) + 6,\ Y(0) = 0 \}$$

The solution now involves integrals of cubic polynomials.

```
> solb:=dsolve(",Y(z)); singb:=series(op(2,solb),z=1,3);
dsolve/diffeq/dsol1:   -> first order, first degree methods :
dsolve/diffeq/dsol1:   trying linear bernoulli
dsolve/diffeq/linearsol:   solving 1st order linear d.e.
dsolve/diffeq/dsol1:   linear bernoulli successful
```

$$solb := Y(z) = \frac{\int_0^z \mathbf{e}^{(1/3\,u\,(2\,u^2+3\,u+6))}\,du\ \mathbf{e}^{(-1/3\,z\,(2\,z^2+3\,z+6))}}{z^2 - 2\,z + 1}$$

$$singb := \int_0^1 \mathbf{e}^{(1/3\,u\,(2\,u^2+3\,u+6))}\,du\ \mathbf{e}^{(-11/3)}\,(-1+z)^{-2} + \left(-6\int_0^1 \mathbf{e}^{(1/3\,u\,(2\,u^2+3\,u+6))}\,du\ \mathbf{e}^{(-11/3)} + \mathbf{e}^{(11/3)}\,\mathbf{e}^{(-11/3)}\right)(-1+z)^{-1} + \left(15\int_0^1 \mathbf{e}^{(1/3\,u\,(2\,u^2+3\,u+6))}\,du\ \mathbf{e}^{(-11/3)} - 3\,\mathbf{e}^{(11/3)}\,\mathbf{e}^{(-11/3)}\right) + O(-1+z)$$

```
> c2:=factor(simplify(coeff(singb,z-1,-2)));
  c1:=factor(simplify(coeff(singb,z-1,-1))); C3:=evalf(c2):
```

$$c2 := \int_0^1 \mathbf{e}^{(1/3\,u\,(2\,u^2+3\,u+6))}\,du\ \mathbf{e}^{(-11/3)}$$

$$c1 := -6\int_0^1 \mathbf{e}^{(1/3\,u\,(2\,u^2+3\,u+6))}\,du\ \mathbf{e}^{(-11/3)} + 1$$

```
> mb1:=c2*(n+1)-c1; evalf(mb1,30);
```

In particular, the mean occupatio ratio is an interesting integral that evaluates to .200973369978844243166574354875...

$$mb1 := \int_0^1 \mathbf{e}^{(1/3\,u\,(2\,u^2+3\,u+6))}\,du\ \mathbf{e}^{(-11/3)}\,(n+1) + 6\int_0^1 \mathbf{e}^{(1/3\,u\,(2\,u^2+3\,u+6))}\,du\ \mathbf{e}^{(-11/3)} - 1$$

$$.200973369978844243166574354875\,n + .40681358985190970216602048413$$

And finally, the approximation is still very good, being within 1% already for $n = 5$, although the asymptotic regime takes a little longer to establish

```
> for j from 0 to 30 by 5 do
  j,evalf(subs(u=1,diff(gb(j,3),u))-subs(n=j,mb1),30) od;
```

0, -.406813589851909702166020484l3

5, -.011680439746130917998892258 50

10, -.003213956307018800498430699 55

15, -.000126982247416062507348649 96

20, $-.5860589380011797239595 65\ 10^{-5}$

25, $-.136929453612279067279 06\ 10^{-6}$

30, $-.22660803401721997292 8\ 10^{-8}$

Conclusions

Our purpose here has been to demonstrate how one naturally arrives at the solution of a probabilistic problem using tools like Gfun. Once the solutions have been "guessed", it is possible to come back, think, and prove solutions. For instance, the problem of the mean leads to recurrences that involve history (summation) and a factor of $\frac{1}{n}$; thus, it is reasonable to expect to be within reach of the theory of holonomic functions on which Gfun is based, and rough bounds on the order of recurrences or differential equatiosn suffice to validate the "guesses". In this way, we have "naturally" rediscovered a solution of the generalized problem due to David Rothman (fatter men) and obtained a variance analysis for the basic problem that appears to be new. The whole session (including the variance computations) takes about 60 seconds of CPU time on a DEC-3000 station.

About the original problem, we may compare the mean seat occuptation to the best possible seating arrangement:

```
> for i from 1 to 3 do `C`.i/(1/(i+1)) od;
```

.8646647168

.8236529631

.8038934796

Thus, we have determined the price to be paid for random access: it is about 15% to 20%. As we saw repeatedly, the asymptotic approximations obtained are extremely good, already for $n = 5$. Also the distributional analysis, where a small variance is obtained, shows that the average-case is highly representative of what will be observed in practice.