Notes on Feynman's Restaurant Problem Notes

by Michael A.Gottlieb

Page 1

F. first observes that if you are dining for only one night, your selection is (uniform) random, so the mean score acquirable is 50 (out of 100).

He goes on to consider the mean score acquirable for exactly two nights of dining, when employing the following strategy: If the rating of the restaurant on the first night is better than average (50/100), go back to it again on the second night; otherwise try a new restaurant on the second night.

For convenience of calculation F. changes the range of the restaurant ratings to the unit interval [0,1].Calling the rating of the first night's restaurant x, that of the second night's (if new) y, and setting $s = ½$ (formerly 50/100), he observes that when $x>s$, the total score will be $2x$, while otherwise it's $x+y$. [Thus when we have already dined at one (or more) restaurants and there is exactly one dinner left, 's' serves as a *threshold value* that determines whether we repeat the [best] restaurant we already tried, or try a new one.] To find the mean score F. sums the integral of $2x$ with respect to x from s to 1, with the integral of $x+y$ with respect to x from 0 to s and with respect to y from 0 to 1,

$$\begin{aligned}&\int_s^1 2x\,dx \;+\; \int_0^1\int_0^s (x+y)\,dx dy\\ &=\left(1-s^2\right)+\;\left(s^2/2+s/2\right)\\ &=1+s/2-s^2/2\\ &=9/8\end{aligned}$$

Next, F. considers a strategy for two nights of dining when we have already tried one or more of the restaurants, of which the best had rating P (referred to hereafter as "the favorite"). Again he calls the rating of the first night's restaurant (if new) x, that of the second night's (if new) y, and he defines a constant P_0, to be determined, such that if $P > P_0$, we forego the possibility of trying any new restaurants and dine at the favorite on both nights. [Thus P_0 serves as a threshold value when *two* nights of dining remain, analogous to $s = ½$ when *one* night remains.] This implies that P_0 should be chosen so that when $P = P_0$, the score for returning to the favorite both nights ($2P$) is the same as the expected score when *not* doing so (which needs to be determined as a function of P). Noting that P_0 must be greater than ½ (the average rating of a randomly selected restaurant), F. goes on to describe his strategy as follows:

If $P \le P_0$, we try a new restaurant (rated x) on the first night, and then on the second night there are three possibilities: try a new restaurant, return to the favorite or return to the first night's restaurant, and the choice is determined by the relative values of P, x and ½, as follows:

$x < P$ and $P <$ ½, or
$x > P$ and $x <$½,

In these two cases the favorite and first night's restaurants are both below average, so we try a new restaurant (rated y) on the second night, scoring $x + y$.

$x < P$ and $P >$ ½,

The favorite is both better than average and better than the first night's restaurant, so we return to the favorite on the second night, scoring $x + P$.

$x > P$ and $x >$½.

The first night's restaurant is both better than average and better than the favorite, so we return to the first night's restaurant on the second night, scoring $2x$.

Finally, if $P > P_0$ then we return to the favorite for *both* nights, scoring $2P$.

To simplify the calculation of the mean score acquirable with the above-described scheme, F. considers the following three cases:

(1) $P_0 < P$

As noted above, the score in this case is $2P$.

(2) $½ < P < P_0$

In this case we score $2x$ if $x > P$, otherwise we score $x + P$. So the mean score is the integral of $2x$ with respect to x from P to 1, plus the integral of $x + P$ with respect to x from 0 to P, which equals $1 + P^2/2$.

$$\begin{aligned} &\int_P^1 2x\,dx \quad + \int_0^P (x+P)\,dx \\ &= (1-P^2) + \left(P^2/2 + P^2\right) \\ &= 1 + P^2/2 \end{aligned}$$

(3) $P < ½$

In this case we score 2*x* if *x* > ½, otherwise we score *x* + *y*. So the mean score is the integral of *2x* with respect to *x* from ½ to 1, plus the integral of *x*+ *y* with respect to *x* from 0 to ½ and with respect to *y* from 0 to 1, which equals 9/8. This is precisely the same calculation made above for the two-night only case.

$$\begin{aligned}&\int_{1/2}^{1} 2x\,dx \quad + \int_0^1\int_0^{1/2}(x+y)\,dxdy \\ &=(1-1/4)+\ (1/8+1/4) \\ &=9/8\end{aligned}$$

[From these considerations it can be seen directly that when the favorite's rating is less than the threshold value of the two-night-only case, *s* = ½, the favorite will never be chosen and thus has no effect at all on the expected score, which is the same as that of the two-night-only case, having no 'favorite' to choose from initially.]

F. goes no to determine P_0 setting the score in case (1), where we dine at the favorite both nights, equal to the expected score in case (2), where there is yet some possibility of returning to the favorite [while ignoring case (3), for which there is no possibility of returning to the favorite], and solving for *P*. This yields the quadratic equation, $1+P_0^2/2-2P_0=0$, which F. solves using conventional methods, finding $P_0=2-\sqrt{2}\approx 0.586$. He checks himself by calculating $P_0^2=6-4\sqrt{2}$, and $2P_0=4-2\sqrt{2}\approx 1.172$, and plugging them into $2P_0=1+P_0^2/2$.

Finally, F. summarizes his strategy as follows:

If $P < .500$	*x*;	if $x < ½$ *y*, else *x*.
If $.586 > P > .500$	*x*;	if *x*>*P* *x*, else *P*.
If $P > .586$	*P;*	*P*.

Page 2

F. now generalizes his strategy for two nights of dining, to *n* nights, when one has already tried one or more of the restaurants, of which the best had rating *P* ("the favorite"). He introduces the notation P_n for the threshold value to which one compares *P* when *n* nights of dining remain.

Having already found $P_1 = ½$, and P_2 (formerly "P_0") $= 2 - \sqrt{2}$, he shows how to find any P_n by direct analogy with his previous computations. He defines his general strategy (with scores shown on the right) to be:

(A) $P > P_n$	repeat *P* for the *n* remaining nights.	*nP*
(B) $P_{n-1} < P < P_n$	try *x*; if *x* > *P* repeat *x;*	*nx*
	else repeat *P*.	*x + (n-1)P*

The score given by rule (A) is self-explanatory. For the first part of (B) when $x > P$, x becomes (the rating of) the *new* favorite, and when in addition $P_{n-1} < P$ then necessarily $x > P_{n-1}$, thus on the *following* night (A) dictates that we will repeat *x* for the remaining *n* – 1 nights, for a total score (including the previous night) of *nx*. Similarly for second part of (B) when $P \geq x$: in this case *P remains* (the rating of) the favorite, and when in addition we have $P_{n-1} < P$, then on the *following* night (A) dictates that we will repeat *P* for the remaining *n* – 1 nights, for a total score (including the previous night) of *x* + (*n*-1)*P*.

[F. no longer bothers to consider the score when $P < P_{n-1}$, because it isn't needed to calculate P_n. To see this we can extend F.'s strategy as follows:

(C) $P \leq P_{n-1}$	try *x*; if $x > P_{n-1}$ repeat *x;*	*nx*
	else try *y*.	*x + y + ...*

The expectation of the first part of (C) equals ½ + the expected score for n – 1 nights of dining when *x* is the favorite and $x > P_{n-1}$, while the expectation of the second part equals ½ + the expected score when dining for *exactly* n – 1 nights. Clearly P_n can not depend on these (as reflected by the fact that *P* does not appear in the scores of this case).]

To calculate P_n F. solves for *P* when the expected score for case (A), *nP,* is set equal to the expected score for case (B). To find the expected score for case (B) he sums [integral of *nx* with respect to *x* for *x* = *P* to 1] with [integral of *x* + (*n*-1)*P* with respect to *x* from 0 to *P*] , yielding $n/2 + ((n-1)/2)P^2$.

$$\int_P^1 nx\,dx \quad + \int_0^P \left(x + (n-1)P\right)dx$$
$$= \frac{n}{2}\left(1 - P^2\right) + \left(P^2/2 + (n-1)P^2\right)$$
$$= \frac{n}{2} + \frac{n-1}{2}P^2$$

Using conventional means F. then solves the quadratic

$$\frac{n}{2}+\frac{n-1}{2}P_n^2 = nP_n$$

finding that $P_n = \sqrt{n}\Big/\left(\sqrt{n}+1\right)$. As an example, F. calculates $P_{10} \approx 0.76$.

Next, considering the case $n = 16$ ($P_{16} = 0.80$), F. expresses his solution in a *completely different way*, inverting it to calculate n from P_n!

He observes that if $P_n = 0.80$, then there is probability 0.80 that a randomly selected restaurant will have a lower rating. How *much* lower? On average, it will be 0.40 lower. So the expected loss is 0.80 x 0.40 = 0.32. On the other hand, there is probability 0.20 that a randomly selected restaurant will have a higher rating, and on average it will be 0.10 higher, and when that happens ($P > P_n$) we sample it n times, so the expected gain is 0.2 x $0.1n = 0.02n$. To solve for n, one sets the expected gain to the expected loss $0.02n = 0.32$, thus $n = 16$.

F. then generalizes his inverted solution, as follows:

Below chance P	average loss $P/2$	expectation $P^2/2$
Above chance $1 - P$	average gain $(1-P)/2$	expectation $\frac{(1-P)^2}{2}n$

By equating the expected loss to the expected gain for the "break even" case $P = P_n$, F. finds

$$n = \frac{P_n^2}{\left(1-P_n\right)^2} \quad \text{or} \quad \sqrt{n} = \frac{P_n}{\left(1-P_n\right)},$$

which is just an inversion of the solution for $P_n = \sqrt{n}\Big/\left(\sqrt{n}+1\right)$.

[This almost looks like *magic*, but I suspect F. found this way of looking at the problem by examination of his integrals for cases (A) and (B) above, which equated yield $P = P_n$. Indeed, F. was *very good* at doing algebra in his head, and probably noticed that

$$\frac{n}{2}+\frac{n-1}{2}P_n^2 = nP_n$$

can be rearranged as

$$\frac{P_n^2}{2} = \frac{\left(1-P_n\right)^2}{2}n,$$

which could have been suggestive, to Feynman's way of thinking.]