The M/G/1 Vacation Paradox and Wald's Identity

Michael J. Neely
http://www-bcf.usc.edu/~mjneely

I. VACATION PARADOX

Consider an M/G/1 queue with i.i.d. vacations $\{V_i\}_{i=1}^{\infty}$. Arrivals are Poisson of rate λ. Suppose we start empty at time 0 and we take back-to-back vacations until we get the first arrival (in which case we finish the last vacation and then start a busy period). Each new vacation we take is i.i.d. and distributed with a general CDF $F_V(v)$ with mean $\mathbb{E}[V]$. Let I be the duration of the idle interval before the first busy period. The idle interval consists of back-to-back vacations:

$$I = V_1 + V_2 + ... + V_K \tag{1}$$

where K is the random number of vacations during the idle interval. The variable V_1 has CDF $F_V(v)$. But do the others? What is the expected size of V_K? Since we know that V_K is the *last* vacation of the idle interval, it means that it must have been long enough for at least one arrival to occur. So we expect that $\mathbb{E}[V_K] \geq \mathbb{E}[V]$, and we expect strict inequality to hold whenever the vacations are not a deterministic constant. Let's call this the *vacation paradox*. In fact, $\mathbb{E}[V_K] = \mathbb{E}[V|V > Y]$, where V is a random variable with CDF $F_V(v)$ and Y is an independent exponential random variable with rate λ. Specifically:

$$\mathbb{E}[V|V > Y] = \int_0^{\infty} v f_{V|V>Y}(v|V > Y)dv = \frac{\int_0^{\infty} ve^{-\lambda v} f_V(v)dv}{\int_0^{\infty} e^{-\lambda v} f_V(v)dv}$$

Another strange observation: We know that V_1 has CDF $F_V(v)$. However, if we are told that $K = 12$, the *conditional* CDF of V_1 changes, since the new information means that V_1 was too small for an arrival to occur.

II. THE RELATIONSHIP BETWEEN $\mathbb{E}[I]$ AND $\mathbb{E}[K]$

A. *Sloppy, clever, and very clever students*

Let K be the random number of vacations during an idle interval. The random variable K is related to I via (1). What is the relationship between $\mathbb{E}[K]$ and $\mathbb{E}[I]$? Let's see how some virtual students might guess the relationship:

- A *sloppy student* guesses $\mathbb{E}[I] = \mathbb{E}[V]\mathbb{E}[K]$ from equation (1), thinking that since all the V_i are i.i.d, we can just take expectations in this way. This student does not try to justify this equation.
- A *clever student* observes that the vacations in equation (1) are *not* all i.i.d., since the last vacation V_K is expected to be larger than average. So the clever student guesses $\mathbb{E}[I] \neq \mathbb{E}[V]\mathbb{E}[K]$.
- A *very clever student* does the calculations:

$$\mathbb{E}[I] = \mathbb{E}[V] + q\mathbb{E}[I] \implies \mathbb{E}[I] = \frac{\mathbb{E}[V]}{1-q}$$
$$\mathbb{E}[K] = 1 + q\mathbb{E}[K] \implies \mathbb{E}[K] = \frac{1}{1-q}$$

  where $q = P$[no arrivals during the first vacation]. So the very clever student concludes $\mathbb{E}[I] = \mathbb{E}[V]\mathbb{E}[K]$.

Thus, the very clever student gives the same answer as the sloppy student!

B. *Student Reactions*

- The sloppy student ignores the clever student and asserts that the equation $\mathbb{E}[I] = \mathbb{E}[V]\mathbb{E}[K]$ is "obvious."
- The very clever student listens to the reasoning of the clever student. This makes the very clever student confused. Together, the clever and very clever students explore the question "why should $\mathbb{E}[I] = \mathbb{E}[V]\mathbb{E}[K]$ be true"? These students agree to explore the "why" question by looking into *Wald's identity*.

C. *Wald to the rescue*

Wald's identity treats random sums of random variables. Let $\{V_i\}_{i=1}^{\infty}$ be i.i.d. variables. For simplicity, assume the V_i variables are nonnegative.[1] Let K be a random *stopping time* with the property that, for each $i \in \{1, 2, 3, ...\}$, the event $\{K \geq i\}$ is independent of V_i. Then:

$$\mathbb{E}\left[\sum_{i=1}^{K} V_i\right] = \mathbb{E}\left[\sum_{i=1}^{\infty} V_i 1_{\{K\geq i\}}\right] = \sum_{i=1}^{\infty} \mathbb{E}\left[V_i 1_{\{K\geq i\}}\right] \overset{(a)}{=} \sum_{i=1}^{\infty} \mathbb{E}[V_i]\mathbb{E}\left[1_{\{K\geq i\}}\right] = \mathbb{E}[V]\sum_{i=1}^{\infty} P[K \geq i] = \mathbb{E}[V]\mathbb{E}[K]$$

where (a) uses the fact that V_i is independent of the event $\{K \geq i\}$. The answer is *simple*, but it is certainly not *obvious*. Wald's identity can be used to compute the expectation of the sum of i.i.d. random variables that stop at the first index K such that the sum of the first K variables exceeds a given threshold θ. It is also used to prove the *elementary renewal theorem*.

[1]One is always allowed to push an expectation through an infinite sum of nonnegative random variables.