Planning the Macy’s Special Sunrise Sale

Your mission: Your mission, if you choose to accept it, is to help Macy’s Department Store plan for a special sunrise sale at its flagship 34th Street store in Manhattan. On the selected day, customers entering the store between 6am and 8am will be eligible for a 50% discount on all their purchases made during the two-hour period after their time of arrival. Upon entering the store, each customer will be given a special hand-held electronic device, on which is indicated the current time and the customer’s time of arrival. Each customer enters into this device product codes and quantities for the items they want to purchase. The times these purchase entries are made are automatically recorded. Each customer is allowed to purchase all the items entered into the electronic device within the two-hour time period, up to some upper dollar limit. (The timing provided by the electronic device allows customers to complete purchases even if the customers are stuck waiting in checkout lines after their time limit has expired. Incidentally, it also specifies order of purchase to allocate limited supplies on the day of purchase. We assume all orders can be met, which might require delays if inventory on hand is depleted.)

Assumptions: Managers estimate that customers will arrive at the store at a rate of 30 per minute throughout the two-hour sale period. They estimate that customers will stay in the store for an average of one hour and spend (after the discount) an average of $400 each. Assume that the arrival process is a Poisson process (starting at 6am). Assume that the times spent in the store by customers are independent and identically distributed (i.i.d.) random variables, independent of the arrival process, each uniformly distributed on the interval [0, 2], measuring time in hours. Assume that the dollars spent in the store (again, after discounts) by customers are i.i.d. random variables, independent of the arrival process, each exponentially distributed with mean $400.00.

1. Grand Totals. (30 points)

(a) What is the expected total number of customers to come to this sale?

Let $N$ be the total number of customers that come to the sale. Then

$$E[N] = 30 \text{ per minute} \times 120 \text{ minutes} = 3600$$

(b) What is the (approximate) probability that the total number of customers to come to this sale is at least 3660?
The total number has a Poisson distribution with mean 3600 from part (a). For a Poisson distribution, the variance equals the mean. Since the standard deviation is the square root of the variance, the standard deviation here is 60. The Poisson distribution is approximately normal, provided that the mean is not too small (greater than equal to 5, say). We see that 3660 is 1 standard deviation above the mean, so that the probability of exceeding 3660 is about 0.16, from the normal table. In more detail, let $N$ be the total number. Then

$$P(N \geq 3660) = P(N - 3600 \geq 3660 - 3600)$$
$$= P\left(\frac{N - 3600}{60} \geq \frac{3660 - 3600}{60}\right)$$
$$= P\left(\frac{N - 3600}{60} \geq 1\right)$$
$$\approx P(N(0, 1) \geq 1)$$
$$= 1 - P(N(0, 1) \leq 1) = 1 - 0.8413 \approx 0.16$$

(c) What is the expected total amount of the proceeds (in dollars, after discounts) from this sale?

Let $T$ be the total proceeds. Then

$$E[T] = 3600 \text{ expected number of arrivals} \times \$400 \text{ per customer} = \$1,440,000$$

(d) What is the standard deviation of the total amount of proceeds from the sale?

The total proceeds is a compound Poisson process, so that the total proceeds has a compound Poisson distribution. The variance of the total proceeds is the mean arrival rate multiplied by the second moment of the proceeds per customer. Since the exponential proceeds from each customer has mean 400, it has variance 160000 and thus second moment 320,000. The total proceeds has variance equal to $3600 \times 320,000$. The standard deviation is then

$$\sigma = 60 \times 400 \times \sqrt{2} = 24000\sqrt{2} \approx 24,000 \times 1.414 \approx 34,000$$

(e) What is the (approximate) probability that the total proceeds is at least $1,406,000 but no more than $1,508,000?

The lower limit is one standard deviation below the mean, while the upper limit is two standard deviations above the mean, so we have $P(-1 \leq N(0, 1) \leq 2) = 0.977 - 0.16 \approx 0.82$ In more detail, let $T$ be the total proceeds. Then

$$P(1,406,000 \leq T \leq 1,508,000)$$
$$= P(1,406,000 - 1,440,000 \leq T - 1,440,000 \leq 1,508,000 - 1,440,000)$$
\[ P \left( \frac{1,406,000 - 1,440,000}{34,000} \leq T - 1,440,000 \leq \frac{1,508,000 - 1,440,000}{34,000} \right) \]
\[ = P \left( -1 \leq \frac{T - 1,440,000}{34,000} \leq 2 \right) \]
\[ \approx P \left( -1 \leq N(0,1) \leq 2 \right) \]
\[ = P \left( N(0,1) \leq 2 \right) - P \left( N(0,1) \leq -1 \right) = 0.9772 - 0.1587 = 0.8185 \]

2. Customers in the Store. (46 points)

Let \( N(t) \) be the number of customers shopping in the store at time \( t \) among those participating in the sale (who have arrived before time \( t \) at an allowed time in the time interval \([6, 8]\), are still eligible to make purchases with the sale discount, and have not yet departed).

(f) What is \( E[N(11)] \), the expected number of customers still shopping in the store at 11am (among those participating in the sale)?

\[ E[N(11)] = 0 \] because the customers participating in the sale all arrive by 8am and are allowed to purchase only for at most two hours.

(g) What is \( E[N(8)] \), the expected number of customers still shopping in the store at 8am (among those participating in the sale)?

We now use the fact that the number \( N(t) \) is the number in an infinite-server queue with a Poisson arrival process. Below we give the relevant figure, following the posted 1993 Physics paper:

We now proceed analytically. Let \( m(t) = E[N(t)] \). But it is convenient to relabel time so that 6am is time 0 and 8am is time 2; that is, we subtract 6 from all times. Then we have

\[ m(t) = \int_0^t \lambda(u)G(t-u) \, du = \lambda \int_0^t 1 - G(t-u) \, du \]

by the 1993 Physics paper or by Example 5.18. In this case, the service-time distribution has been assumed to be uniform on the interval \([0, 2]\), so that \( G(t) = t/2, 0 \leq t \leq 2 \), with density \( g(t) = 1/2, 0 \leq t \leq 2 \). Hence, (letting 6am be time 0 and letting 8am be time 2) we have

\[ m(2) = 1800 \int_0^2 (1 - (t/2)) \, dt = 1800 \]

(The expected total number of arrivals is 3600. Half of these are expected to be there at time 8am. The area of the triangle is half the total area.)

(h) What is the variance of \( N(8) \)?
Since $N(8)$ has a Poisson distribution, the variance equals the mean. Hence $\text{Var}(N(8)) = E[N(8)] = 1800$.

(i) What is $E[N(9)]$, the expected number of customers still shopping in the store at 9am, and are still eligible to participate in the sale, among those who arrived between 6am and 8am?

We now give the new Poisson random measure view: The area is now 1/4 of what it was before. Some of the arrivals between 7am and 8am will still be in the system at time 9am and still be eligible to purchase under the sale. Using the time shift by 6 introduced above, we have

$$m(3) = \lambda \int_{1}^{2} (1 - (3 - t)/2)) \, dt = 1800 \times \frac{1}{4} = 450$$

(j) Does $N(t) - N(s)$ have a Poisson distribution for each pair of times $(s, t)$ with $s < t$? Explain.

No, there is nothing to prevent $N(s)$ being greater than $N(t)$, which allows negative values. The Poisson distribution concentrates on the nonnegative integers. To elaborate (which is not really necessary for you to do), $N(t) - N(s)$ can be seen to be the difference of two independent Poisson random variables: $X(s, t)$ - the number arriving after time $s$ and before time $t$ that are still in the system at time $t$ - minus $Y(s, t)$ - the number arriving before time $s$ that depart between times $s$ and $t$ (and thus are no longer in the system at time $t$). These can be seen
to be independent Poisson random variables. This difference could be negative, which is not possible for a Poisson distribution.

(k) Is \( \{N(t) : t \geq 0\} \) a nonhomogeneous Poisson process? Explain.

No, a nonhomogeneous Poisson process has independent Poisson increments. We have just seen that the increments could be negative, and so they cannot be Poisson. To elaborate (which again is not really necessary), the process fails to have independent increments. Consider \( t_1 < t_2 < t_3 < t_4 \). The number that arrived in the interval \([t_1, t_2]\) but are still in the system and depart in the time interval \([t_3, t_4]\) are counted in both random variables. These events are counted positively in \( N(t_2) - N(t_1) \), but negatively in \( N(t_4) - N(t_3) \).

(m) Is \( \{N(t) : t \geq 0\} \) a Markov process? Explain.

No, because the probability of future events given the present state and past history does NOT only depend on the present state. The arrival process presents no problem, because it is a Poisson process. That makes future arrivals independent of the present state \( N(t) \) as well as the history. The problem occurs because the service-time distribution has a uniform distribution. Since it is not exponential, it fails to have the lack-of-memory property. From the entire history, we get extra information about the likelihood of future departures. If a customer has been in service for a long time (relative to the possibilities in \([0, 2]\), then a departure soon is more likely. From the full history, we know how long each customer in service has been in
service. So the history gives us extra information beyond the present state, as far as describing the future is concerned.

3. Random Variables from a Random Sample. (24 points)

Suppose that 5 customers who have participated in the sale are selected at random as they are leaving the store. Let \( X_i \) be the total value of customer \( i \)'s purchases, \( 1 \leq i \leq 5 \). Specify the mean value and the probability density function for each of the following random variables

(m) the sum of their purchases: \( S \equiv X_1 + \cdots + X_5 \)

Since the random variables \( X_i \) are exponential, the random variable \( S \) has a gamma distribution; see Section 5.23 and Example 2.38. In particular,

\[
f_S(t) = \lambda e^{-\lambda t} \frac{(\lambda t)^4}{4!}, \quad t \geq 0.
\]

where \( 1/\lambda = E[X_1] = 400 \). The expectation of the sum is the sum of the expectations, so that

\[
\]

(n) the minimum value of their purchases: \( Y \equiv \min \{X_1, \ldots , X_5\} \)

The minimum of i.i.d. exponentials is again exponential with a rate equal to the sum of the rates; see the Concise Summary. Thus

\[
f_Y(t) = 5\lambda e^{-5\lambda t}, \quad t \geq 0.
\]

where \( 1/\lambda = E[X_1] = 400 \). The expectation is

\[
E[Y] = \frac{1}{5\lambda} = \frac{400}{5} = 80.
\]

(o) the maximum value of their purchases: \( Z \equiv \max \{X_1, \ldots , X_5\} \)

The maximum is a bit more complicated, but it is not too difficult if we start by computing the cdf:

\[
F_Z(t) = P(Z \leq t) = P(X_1 \leq t, \ldots , X_5 \leq t) = (1 - e^{-\lambda t})^5, \quad t \geq 0.
\]

where \( 1/\lambda = E[X_1] = 400 \). We can then differentiate to get the probability density function (pdf):

\[
f_Z(t) = 5(1 - e^{-\lambda t})^4\lambda e^{-\lambda t}, \quad t \geq 0.
\]

We can then expand the power to get a form from which we can compute the mean.

To compute the mean of the maximum, it is convenient to represent the maximum as a sum of independent minima. At this point, it is perhaps helpful to think of 5 people entering an empty post office with many clerks (servers), so that all five can enter service at once. Our
question is equivalent to: When does the last person finish service. This is easy to analyze when the service times are exponential. The time that the first person finishes is the minimum of the 5 exponential random variables. After that first person finishes, there are 4 people left. By the lack-of-memory property, their remaining service times are new i.i.d. exponentials, and so forth.

As a consequence of this reasoning, the maximum value of $k$ i.i.d. exponentials is the minimum value plus the maximum of $k-1$ i.i.d. exponentials. We would first find the minimum. Then, by the lack-of-memory property of the exponential distribution, the remaining $k - 1$ values would be distributed as new i.i.d. exponentials with the same mean. We can then continue recursively to get

$$max\{X_1, X_2, X_3, X_4, X_5\} \overset{d}{=} min\{X_1, X_2, X_3, X_4, X_5\} + min\{X_1, X_2, X_3, X_4\} + \cdots + min\{X_1, X_2\} + X_1 ,$$

where $\overset{d}{=} \text{ means equal in distribution and the minimum random variables on the right are understood to be independent. This is useful, because the minimum of i.i.d. exponential random variables is itself exponential with a rate equal to the sum of the rates (and thus a mean equal to the individual mean divided by the number of component exponential random variables). Hence,}

$$E[max\{X_1, X_2, X_3, X_4, X_5\}] = E[min\{X_1, X_2, X_3, X_4, X_5\}] + E[min\{X_1, X_2, X_3, X_4\}] + E[min\{X_1, X_2, X_3\}] + E[min\{X_1, X_2\}] + E[X_1],$$

$$= \frac{1}{5\lambda} + \frac{1}{4\lambda} + \frac{1}{3\lambda} + \frac{1}{2\lambda} + \frac{1}{\lambda},$$

$$= (400) \left( \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} \right),$$

$$= (400) \frac{137}{60} = 913.333.$$

---