# Queueing and Gap Acceptance

CE 391F

April 16, 2013

# REVIEW

What kinds of distributions are we considering in this class?

How do we generate a $U(0,1)$ number?

# OUTLINE

# Generating other kinds of random number

1. Uniform between $A$ and $B$
2. Uniform integers
3. Generic distributions
4. Exponential distribution
5. Standard normal distribution
6. Nonstandard normal distribution

# INTERSECTIONS AND QUEUEING

The last topic for the course is intersections where one vehicle stream must yield to another.

1. Two-way stops
2. Permitted turns at signals
3. Roundabouts

Protected turns or nonconflicting movements at signals are less interesting from the standpoint of traffic flow theory, and are more of a traffic operations concern.

(For instance, in CTM you can simply set capacity to zero during the red interval, or in CA add a rule preventing a vehicle from passing a cell.)

Assume that we have two traffic streams:

Priority stream: This traffic stream has priority at the junction and essentially operates independently of the other stream.

Minor stream: This traffic stream has to yield to the priority stream, and can only move during gaps.

The two questions we'll ask: how does the size of the gaps relate to the number of vehicles that can move ("gap acceptance"), and how much delay is caused to vehicles ("queuing")?

# GAP ACCEPTANCE

Let's say there is a 10 second gap between vehicles on the priority stream. How many vehicles (if any) will be able to move on the minor stream?

The *critical time* $t_c$ is the minimum gap needed for a vehicle to move.

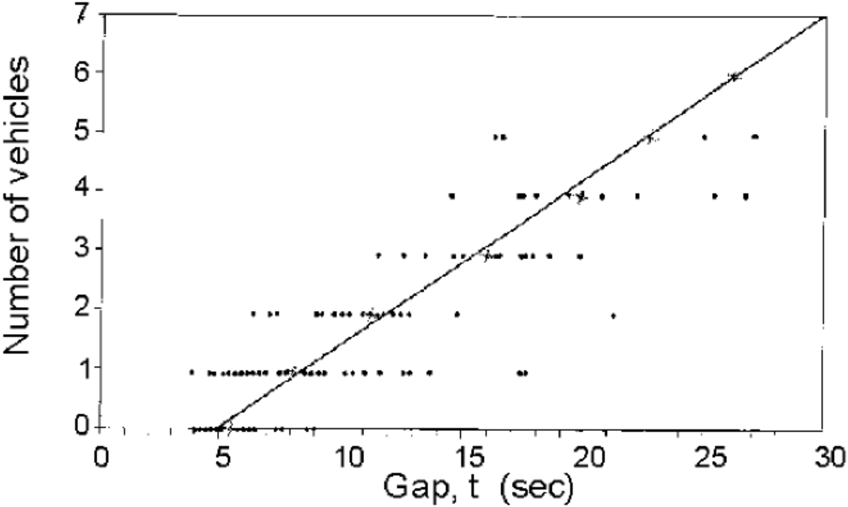
The *follow-up time* $t_f$ is the additional gap needed for vehicles beyond the first to move.

It is hard to specify single values for these parameters, which depend on drivers, roadway geometry, the type of turning movement (right turn vs. left turn), and so forth.

For a particular situation, they can be identified experimentally:

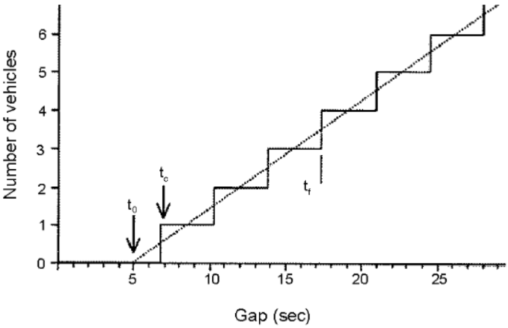
1. Observe the intersection during a period of time in which there is always at least one vehicle waiting from the minor stream.
2. For each gap, record the length of the gap and the number of vehicles which moved.
3. Perform a regression.

The first step is the most important.


Number of vehicles
Gap, t (sec)
0
1
2
3
4
5
6
7
0
5
10
15
20
25
30

For this data, the intercept is $t_0 = 5$ seconds and the slope is $t_f = 3.5$ seconds. The critical gap can be estimated as $t_c = t_0 + t_f/2 = 6.8$ seconds.

(Why not $t_0 + t_f$?)



# Headway distributions

The gaps between vehicles are stochastic. What distributions can be used to model them?

The exponential distribution is common for representing the time between Poisson (*independent*) events.

Recall that the exponential distribution with mean $\lambda$ has pdf $(1/\lambda)\exp(-t/\lambda)$ for $t \geq 0$, and cdf $1 - \exp(-t/\lambda)$.

If the flow rate on the priority stream is $q_p$, then $\lambda = 1/q_p$ and the pdf is $q_p \exp(-tq_p)$.

The independence assumption is fine at low flow rates. At higher flow rates, other distributions are better.

One option is the *displaced exponential* distribution which has a lower bound of $t_m$ (a "minimum" headway)

If we simply translate the cdf $t_m$ units to the right $1 - \exp(-(t - t_m)/\lambda)$ for $t \geq t_m$,, we also need to adjust $\lambda$ so that the mean gap is still $1/q_p$.

$$t_m + \lambda = 1/q_p \iff \lambda = (1 - t_m q_p)/q_p$$

so the cdf is

$$1 - \exp(-(tq_p)/(1 - t_m q_p))$$

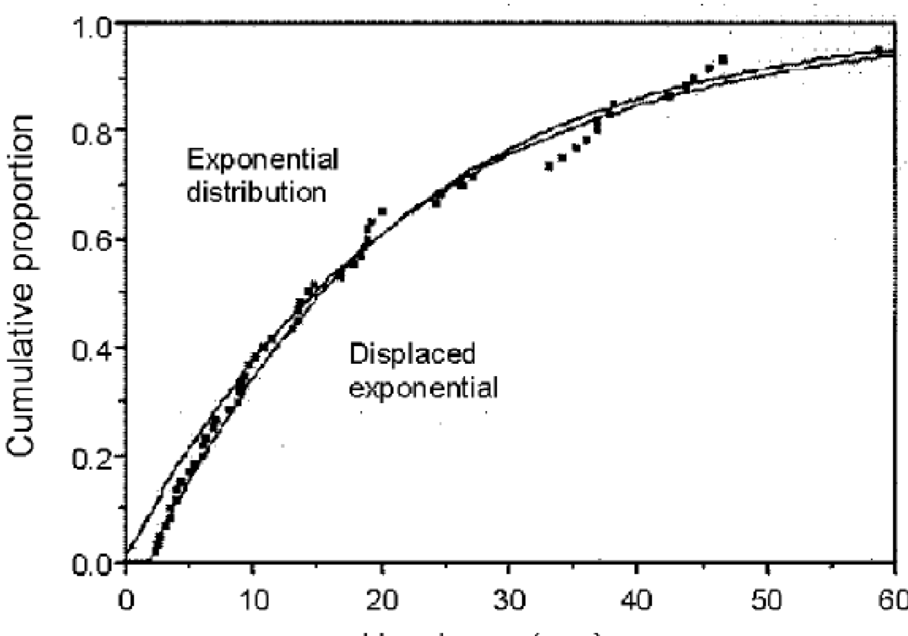
Perhaps a still better option is a dichotomized headway distribution which can account for platooning.

Assume there are two types of vehicles: *bunched* vehicles (whose motion is dictated by car-following) and *free* vehicles (which can travel at their desired speed).

If the proportion of free vehicles is $\alpha$, we can formulate a "combined" distribution where the free vehicles' headways are exponential, and the bunched vehicles' headways are displaced exponential.

The resulting cdf is

$$1 - \alpha \exp(-tq_p) - (1 - \alpha) \exp(-(tq_p)/(1 - t_m q_p))$$


Cumulative proportion
1.0
0.8
0.6
0.4
0.2
0.0
Exponential distribution
Displaced exponential
0
10
20
30
40
50
60

You should know how to sample random numbers from any of these distributions...

# Minor stream capacity

How can we calculate the capacity for the minor stream?

Assume that the pdf for gap size on the priority stream is given by $f(t)$. Let $g(t)$ denote the number of vehicles from the minor stream which can move for a gap of size $t$.

If we are willing to accept fractional $g$ values (interpreted as long-run averages), the linear relationship gives $g(t) = [(t - t_0)/t_f]^+$ where $[\cdot]^+ = \max\{\cdot, 0\}$.

(You will get slightly different results if you round to integer values.)

Then the expected number of vehicles which can move during a single gap is

$$\int_0^\infty g(t)f(t)\ dt$$

.

Assuming the observation period has unit length, there will be $q_p$ gaps in the priority stream.

Thus $q_m = q_p \int_0^\infty g(t)f(t)\ dt$

If the priority stream has exponential headways and the linear relationship holds, then we get

$$q_m = \frac{\exp(-q_p t_0)}{t_f}$$

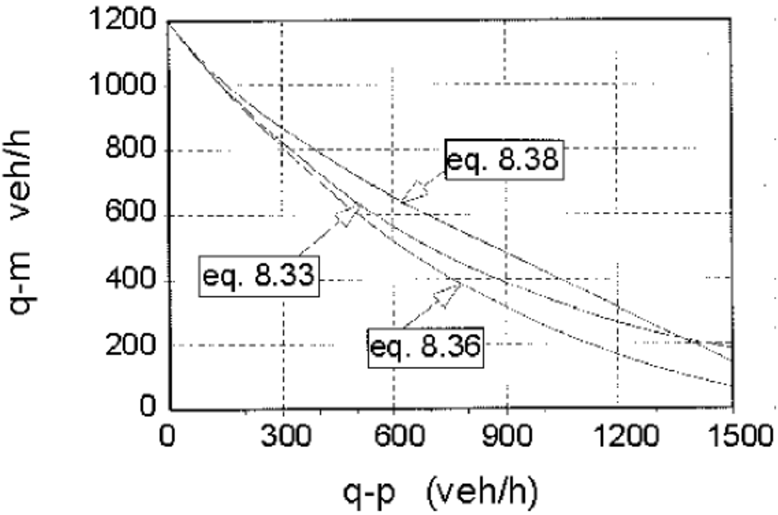
using a common integral "trick" for the exponential distribution.

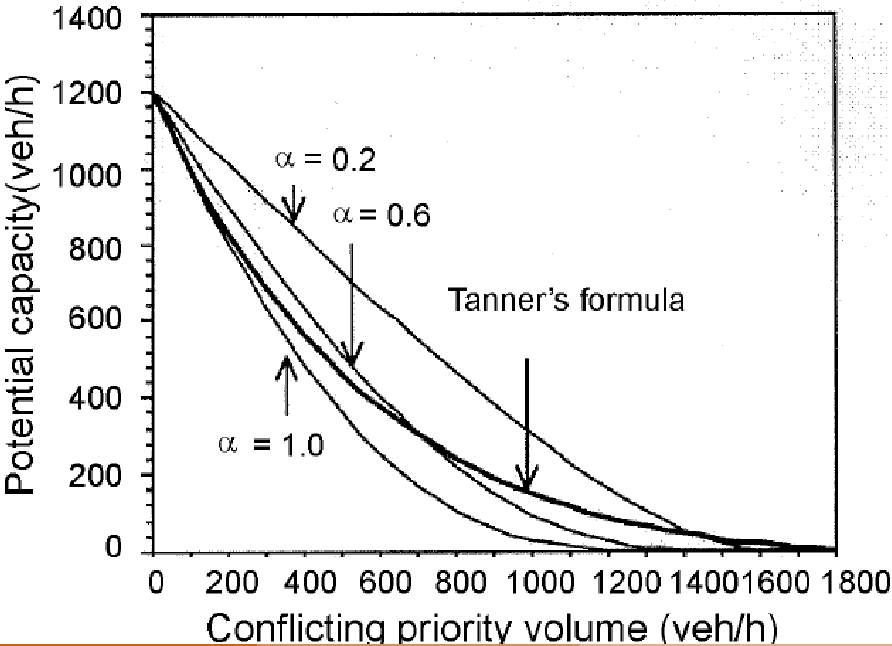
Notice the assumptions that went into this analysis:

- $t_c$ and $t_f$ are constant across the population
- Exponentially-distributed priority stream headways
- Constant traffic volumes for both stream

Relaxing these assumptions leads to slightly different equations.

The equation we derived was 8.33:





Potential capacity(veh/h)
1400
1200
1000
800
600
400
200
0
0
200
400
600
800
1000
1200
1400
1600
1800
Conflicting priority volume (veh/h)
α = 0.2
α = 0.6
Tanner's formula
α = 1.0

# Simulation

However, even these “relaxations” still make other assumptions. Another approach is event-based simulation.

This is different from the simulations we’ve seen so far, which update the state of the system at regular intervals. An *event-based simulation* updates the state when particular events happen.

For instance:

- Another vehicle arrives at the back of the queue.
- A new gap opens in the priority stream.
- A vehicle leaves the queue.

Example

From this kind of simulation it is easy to get measures of effectiveness such as

- Average and maximum queue lengths
- Average and maximum delay to drivers
- Standard deviations, higher-order moments, percentiles, etc. of these quantities
- Capacity for minor stream
- etc. etc.

# QUEUEING THEORY

Queueing theory can develop exact expressions for certain types of queuing situations.

Let $\lambda$ be the average arrival rate (approach volume on the minor stream), and $\mu$ the average departure rate when there are vehicles waiting in queue.

Let $\rho = \lambda/\mu$ be the *intensity* of the queue.

Queueing theory classifies scenarios with notation like M/M/1 or M/D/2.

The first letter indicates how *arrivals* occur. M = Poisson arrivals (exponential headway), D = deterministic, G = general, etc.

The second letter indicates how *departures* occur.

The number indicates the number of service channels. Typically 1 for conflicting traffic movements, but can be higher when applied to other types of queues (toll booths, checkout lines, etc.)

Deterministic queues are not particularly interesting.

Given a D/D/1 queue with initial queue length $L(0)$, the length of the queue at time $t$ is simply $L(t) = [L(0) + (\lambda - \mu)t]^+ = [L(0) + (\rho - 1)\mu t]^+$.

Consider a M/M/1 queue, with Poisson arrivals and departures (exponential headways).

If there are $L > 0$ vehicles in queue, then after a single time step the queue will grow to length $L + 1$ with probability proportional to $\lambda$, and shrink to $L - 1$ with probability proportional to $\mu$.

(Assume that the time step is small enough that only one such action will happen. We can do this and take limits without any problem.)

This gives us the infinite *transition matrix*

$$\begin{bmatrix} 1-\lambda & \mu & 0 & 0 & \\ \lambda & 1-\mu-\lambda & 0 & 0 & \cdots \\ 0 & \lambda & 1-\mu-\lambda & -\lambda & \\ 0 & 0 & \lambda & 1-\mu-\lambda & \\ & \vdots & & & \ddots \end{bmatrix}$$

where the entry in row $i$ and column $j$ gives the probability that the system will transition from state $j$ to state $i$.

This is an instance of a Markov process. If the probability distribution for one time instance is given by the vector $\mathbf{p}(t)$ and the transition matrix is $\mathbf{T}$, then $\mathbf{p}(t+1) = \mathbf{T}\mathbf{p}(t)$.

The *steady-state* conditions can be identified with the eigenvectors of the transition matrix.

If $p_k$ is the steady-state probability of having $k$ vehicles in queue, we can verify that $p_k = (1-\rho)\rho^k$ is an eigenvector of $\mathbf{T}$ whenever $\rho < 1$.

If $\rho \geq 1$, there is no steady-state — the queue will grow without bound for as long as demand exceeds capacity.

Example

From the distribution for $p_k$, we can calculate other quantities of interest:

The mean queue length $L$ is $\sum_{k=0}^{\infty} k p_k = \rho/(1-\rho)$.

Similarly, the standard deviation of the queue length is $\sqrt{\rho}/(1-\rho)$.

Let's say you enter the queue when there are $k$ vehicles in front of you. How long will you have to wait until you are at the head of the queue?

The time before each vehicle leaves is exponentially distributed, so your waiting time will be the sum of $k$ exponential distributions, each with mean $1/\mu$.

The sum of exponentially-distributed random variables has an Erlang distribution, and its mean will be $k/\mu$.

So, the expected waiting time is then

$$\sum_{k=0}^{\infty}(p_k)(k/\mu) = \frac{1}{\mu}\sum_{k=0}^{\infty} k p_k = \frac{1}{\mu}\frac{\rho}{1-\rho} = \frac{\rho}{\mu-\lambda}$$

Once you reach the head of the queue, the mean time before you can find a gap of the right size is $1/\mu$, so your total waiting time has mean $1/(\mu-\lambda)$.

In summary:

The average number of vehicles in queue is $\rho/(1-\rho) = \lambda/(\mu - \lambda)$

The average waiting time for a vehicle in queue is $1/(\mu - \lambda)$

The average number of vehicles in queue is the average waiting time multiplied by the arrival rate.

Does this seem intuitive?

This result is known as *Little's Law*, and holds for any queueing process regardless of arrival or departure type.

The M/G/1 queue provides a more realistic setting, since departures are not Poisson but must fit into available gaps.

This queue is a bit harder to analyze; the key result is the Pollaczek-Khinkine formula which gives the mean queue length as

$$L = \rho + \frac{C\rho^2}{(1-\rho)}$$

where $C = 0.5\left[1 + (\sigma(\mu)/\mu)^2\right]$ where $\sigma(\mu)$ is the standard deviation of the departure rate. (Note that $\sigma(\mu)/\mu$ is the coefficient of variation.)

The waiting time is then immediate from Little's Law (just multiply $L$ by $\lambda$)

The main challenge is calculating $C$.

For Poisson arrivals and departures, $C = 1$

For Poisson arrivals and deterministic departures, $C = 1/2$

For deterministic arrivals and departures, $C = 0$.

# TRANSFORMATION TO TIME COORDINATES

The fact that the steady-state formulas break down when $\rho \geq 1$ is somewhat annoying.

In practice, we *do* see oversaturated conditions... they just don't last forever.

There is a clever mathematical trick for dealing with this.

Remember the (boring) D/D/1 queue with

$$L(t) = [(\rho - 1)\mu t + L(0)]^+$$

Notice that this formula is well-defined no matter what $\rho$ is, because it explicitly captures the time-dependent nature of the queue length. By contrast, the steady-state M/M/1 formulas only held as $t \to \infty$.

Even the deterministic formula predicts $L(t) \to \infty$ when $\rho > 1$, but it provides a meaningful result for finite time as well. Trying to provide meaningful results for finite time with stochastic queues gets pretty ugly.

The coordinate transformation technique essentially changes the steady-state formula so its asymptote is the *deterministic formula*, rather than $\rho = 1$.

Let $\rho_s$ be the steady-state intensity and $\rho_t$ the transformed intensity.

We want $1 - \rho_s = \rho_d - \rho_t$, where $\rho_d$ is the equivalent deterministic intensity.

Since $L(t) = (\rho_d - 1)\mu t + L(0)$, we have $\rho_d = (L(t) - L(0))/\mu t + 1$.

Therefore $\rho_s = 1 - \rho_d + \rho_t = \rho_t - (L(t) - L(0))/\mu t$.

If $C = 1$, then $L(t) = \rho_s/(1 - \rho_s)$