Statistical Models Ch 5 (6 in 2nd ed)

Terminology

- Probability distribution $P(x)$
- Probability mass function (pmf, same thing)
- Probability density function (pdf)
- Cumulative distribution function $F(x)$ (cdf)
- $F(x)$ nondecreasing
- Expectation value $E[g(x)] = \langle g(x) \rangle$ is
  - $\sum g(x)P(x)$ or
  - $\int g(x)P(x)$
Terminology

- Mean \( \mu = \langle x \rangle \)
- \( n^{th} \) moment \( \langle x^n \rangle \)
- Variance \( V(x) = \text{var}(x) = \sigma^2 = \langle (x - \langle x \rangle)^2 \rangle \)
- \( \sigma^2 = \langle x^2 \rangle - \langle x \rangle^2 \) [prove on board]
- Standard deviation \( \sigma = \sqrt{\text{var}} \)
- Mode is \( x \) where \( P(x) \) is maximal
- Median is \( x \) s.t. \( F(x) = 1/2 \)

Central Limit Theorem (nib)

- Say \( P(x) \) has mean \( \mu \) and variance \( \sigma^2 \)
- \( Y \) is the mean of \( N > 30 \) random \( x \) values, then \( P(y) \) is approximately normal

\[
N(x, \mu, \sigma^2) = e^{-(x-\mu)^2/2\sigma^2} / \sqrt{2\pi\sigma}
\]

- \( P(y) = N(y, \mu, \sigma^2/N) = \ldots \)
Uniform Distributions

- **Uniform distribution**
- \( P(x) = \frac{1}{(b-a)} \) on \([a,b]\), 0 elsewhere
- Obtain from uniform distr. \( r \) on \([0,1]\)
  \[ x = (b-a)*r+a \]
- \( \mu = \frac{(a+b)}{2} \)
- \( \sigma^2 = \frac{(b-a)^2}{12} \) [show on board]
- \( F(x) = \frac{(x-a)}{(b-a)} \) on \([a,b]\), 0 for \( x<a \), 1 for \( x>b \)
- Used when you know just limits

Poisson Process

- Events happen at times \( t_i \)
- Independent and memoryless
- For every time interval \([t, t+dt]\) the chance of the event happening is the same (\(dt\) is fixed, but \(t\) is variable)
- The past has no influence over the future
- Described by exponential and Poisson distributions
Poisson Processes in real world

- Winning the lottery if you buy every week
- Meteor strike
- Radioactive decay of nucleus
- Network failures
- Customer arrivals
- Water sound model
  – [www.cs.ubc.ca/~kvdoel/icad04](http://www.cs.ubc.ca/~kvdoel/icad04)

Poisson Process

- Let chance that event happens in small time interval \( dt \) be \( \lambda \cdot dt \). \( \lambda \cdot dt \) is very small for very small \( dt \). (Chance of being struck in a given year (\( dt = 1 \) year) by a major meteorite for example.)

- What is chance \( Q(t) \) that no event has taken place in time interval \([0 \ t]\)? Say \( t = N \cdot dt \), with \( N \) large and \( dt \) small.
Poisson Process

- Think \([0 \, t] = [0 \, dt \, 2dt \, \ldots \, (N-1)dt \, Ndt]\)
- Chance no event in \(0\)-\(dt\) is \((1 - \lambda dt)\)
- Chance of no event in \(0\)-\(2dt\) is \((1 - \lambda dt)^2\)
- \(Q(t) = (1 - \lambda dt)^N = (1 - \lambda dt)^{t/dt} \to e^{-\lambda t}\)
- \(P(t)\) is pdf for event; related to \(Q\) by
  \[Q(t) = 1 - \int_0^t P(x)dx\]
  \[P(t) = e^{-\lambda t} \mu = \sigma = 1/\lambda, \text{ median}=\ln(2)/\lambda\]
- [Derive on board]

Poisson Process

- Let \(N(t)\) number of events at time \(t\)
- What is pmf for \(n\) events after time \(t\)?
  \[P[N(t)=n] = e^{-\lambda t}(\lambda t)^n/n! \text{ (book notation)}\]
  \[P_n(t) = e^{-\lambda t}(\lambda t)^n/n! \text{ (my notation)}\]
- [Prove on board]
- \(<n>= \lambda t \quad \text{Var}= \lambda t\)
- \text{Poisson distribution} \(P(n)\)
- \((\text{Discrete and function of } n)\)
Exponential Distribution

- \( P(t) = e^{-\lambda t} \), \( P(t)/\lambda \) is chance event has not happened yet at time \( t \)
- \( P(t) \) is pdf of event happening at time \( t \)
- \( P(t) \) is pdf of intervals between events
- \( P_0(t) = P(t)/\lambda \)
- These are confusible!
- Exponential processes often counterintuitive, so beware!

Poisson Process/Exp. Distr.

- Machine failure problem revisited (nib)
- See sim5b.m for MATLAB code
- With an exponential distribution with the same mean as the distribution used before the result changes: fixing all parts now loses money
Poisson Process/Exp. Distr.

- Lightbulb problem (nib)
  - 2 bulbs have exponentially distributed life
  - Minimize number of changes per year (if I get up there and change 2 bulbs that counts as 1 change)
    - Change only broken one?
    - Change both if one breaks?
    - See MATLAB code bulbs.m
Poisson Process/Exp. Distr.

- Exercises Ch5 (6 in 2\textsuperscript{nd} ed).
  - 15
  - 19
  - 21 (see website for sol.)
  - 27 (see website for sol.)

Gamma/Erlang

- Pdf of sum of $\beta$ exponentials
- $X = X_1 + \ldots + X_\beta$
- $X_i$ exponentially distributed with $\lambda = \beta \theta$
- $P(t) = \beta \theta (\beta \theta t)^{(\beta-1)} e^{-\beta \theta t} / \Gamma(\beta)$
  - $<t> = 1/\theta$
  - Var(t) = $1/\beta \theta^2$
  - $\Gamma(\beta) = (\beta-1)!$
- Redo ex. 21b with Erlang
What to Remember

- Basic concepts and terms
- Uniform distribution
- Understand Poisson process
  - Derivations on handout
  - Exponential distribution
  - Poisson distribution
- Gamma/Erlang distribution
- Central limit theorem
  - Normal distribution