CSCE 619-600 Networks and Distributed Processing Spring 2017

Review of Probability IV

Dmitri Loguinov
Texas A&M University

January 31, 2017

Agenda

- Expectations (cont'd) and moments
  - Jensen's inequality
  - Higher moments, variance, covariance
  - Correlation and dependence
- Conditional expectations
- Sums of variables and convolution

Expectations 7

- In general, $E[u(X)] \neq u(E[X])$
  - We can, however, infer some properties of $E[u(X)]$
- Theorem: Jensen's inequality
  - For every concave function u:

$$u''(x) \leq 0 \Rightarrow E[u(X)] \leq u(E[X])$$

  - and for every convex function u:

$$u''(x) \geq 0 \Rightarrow E[u(X)] \geq u(E[X])$$

- Previous lecture: X was uniform in $[0,s]$
  - We obtained $E[e^X] = 29.48 \geq 12.18 = e^{E[X]}$
  - The same inequality holds for any distribution of X

Expectations 8

- Notation: sometimes convenient to use this form:

$$E[X] = \int_{-\infty}^{\infty} x dF(x)$$

- Example: $X \sim F(x)$, what is $E[F^2(X)]$?
- Sometimes, direct computation of $E[X]$ is tedious
  - As is the case with the binomial distribution:

$$E[X] = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k}$$

- Re-write X as a sum of n Bernoulli random variables:

$$E[X] = E[\sum_{i=1}^{n} Y_i] = \sum_{i=1}^{n} E[Y_i] = np$$

Expectations 9

- Higher moments
  - The expectation is called the first moment
  - The r-th moment is defined as:

$$E[X^r] = \int_{-\infty}^{\infty} x^r dF(x)$$

- Definition: variance
  - Is a quadratic measure of how far X deviates from its mean $E[X]$:

$$Var[X] = E[(X - E[X])^2] = E[X^2] - E^2[X]$$

Expectations 10

- For independent X,Y variance is linear, i.e., $Var[X+Y] = Var[X] + Var[Y]$
- Definition: standard deviation
  - Square root of $Var[X]$:

$$\sigma = \sqrt{Var[X]}$$

- You should know how to derive the mean and variance of the following distributions (Wolff p. 21)
  - Binomial: mean np, variance npq
  - Geometric: $(1-p) / p$ and $(1-p) / p^2$
  - Uniform: $(b + a) / 2$ and $(b - a)^2 / 12$
  - Exponential: $1 / \lambda$ and $1 / \lambda^2$

Expectations 11

- For completeness, we mention two more distributions:
  - Gaussian with PDF:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$$

  - Poisson with PMF:

$$P(X = i) = \frac{e^{-\lambda}\lambda^i}{i!}$$

- Task: derive the mean of the Poisson distribution

Expectations 12

- Hint: use the following equality

$$e^x = \sum_{i=0}^{\infty} \frac{x^i}{i!}$$

  - and notice:

$$E[X] = e^{-\lambda} \sum_{i=0}^{\infty} \frac{i\lambda^i}{i!} = e^{-\lambda} \sum_{i=1}^{\infty} \frac{\lambda^i}{(i-1)!}$$

Expectations 13

- Correlation between variables means when one is large (small), the other tends to be large (small)
- Definition: Two variables are called correlated if and only if their covariance is non-zero:

$$Cov[X,Y] = E[(X-E[X])(Y-E[Y])] = E[XY]-E[X]E[Y]$$

- Example
  - Suppose X is uniform in $[0,1]$ and $Y = X + 2$
  - Clearly, they are dependent, but are they correlated?

$$E[XY] = \int_0^1 \int_2^3 xyf(x,y)dydx$$

Expectations 14

- The difficulty is in computing $f(x,y)$
  - Simpler to use another approach:

$$W = XY = X(X+2) \Rightarrow$$
$$E[W] = \int_0^1 x(x+2)f(x)dx = 4/3$$

  - Then:

$$Cov[X,Y] = 4/3 - 1/2 \cdot 5/2 = 0.0833$$

- Theorem: independent variables are uncorrelated

$$Cov[X,Y] = E[XY] - E[X]E[Y] = 0$$

- Corollary: Correlated variables are always dependent

Expectations 15

- Thus, there are 3 possibilities:
  - Independent and uncorrelated
  - Dependent and correlated
  - Dependent and uncorrelated
- Example of dependent, but uncorrelated variables?
  - X is uniform in $[-1, 1]$ and $Y = |X|$
  - Notice that 50% of the time $X - E[X]$ and $Y - E[Y]$ match in sign, while the other 50% they do not
- Proof that covariance is zero
  - Define $W = XY$, then $W = X^2$ with probability ½ and $W = -X^2$ also with probability ½
  - Thus $E[W]$ = 0 and:

$$Cov[X,Y] = E[W] - E[X]E[Y] = 0 - 0 = 0$$

Conditional Expectations

- Definition: the conditional expectation of a random variable X given some event A with $P(A)>0$ is:

$$E[X|A] = \sum_{i=0}^{\infty} iP(X = i|A)$$

$$E[X|A] = \int_{-\infty}^{\infty} x \cdot dP(X \leq x|A)$$

conditional density

- Example: $X \sim \exp(\lambda)$, obtain $E[X \mid X<10]$
- Prove at home:

$$E[X] = \sum_{y=0}^{\infty} E[X|Y = y]P(Y = y)$$

$$E[X] = \int_{0}^{\infty} E[X|Y = y]f_Y(y)dy$$

- Hint: swap the order of summations/integrals

Conditional Expectations 2

- For non-negative variables

$$E[X|A] = \int_0^\infty P(X > x|A)dx$$

- Quiz preparation:
  - Compute $E[1/X]$ for uniform X in $[1,2]$; what about uniform in $[-2,3]$?
  - Compare $E[1/(X+1)]$ and $1/(E[X]+1)$ for exponential X
  - Define $Z = X^{0.5}$, where X is a non-negative variable with CDF $F(x)$. Express the PDF of Z using $f(x) = F'(x)$.
  - Prove independence of A^c and B^c given than A and B are
  - Derive the following assuming X is defined in $[0,\infty)$

$$P(A) = \int_0^\infty P(A|X = x)f_X(x)dx$$

Example

- Randomized black-jack game
  - Decide on a number a and write it down on paper
  - Then the dealer draws a random number X
  - If $X+a$ is above 10, you bust and win $0
  - If $X+a$ is below 10, you win $X+a$ dollars
- What's the best number to write down when X is exponential with mean 5?
- If you have to pay $\$4.60$ per round, what's your expected gain/loss after n rounds?

Example 2

- Assign function $w(a,X)$ to represent your winnings given that you wrote a and the dealer drew X (w is a function of a random variable)
- Then we need to compute $E[w(a,X)]$
  - Define:

$$w(a,X) = \begin{cases} a+X & a+X \leq 10 \\ 0 & otherwise \end{cases}$$

  - Then:

$$E[w(a,X)] = \int_0^\infty w(a,x)f(x)dx$$

Example 3

- For exponential X this becomes:

$$E[w(a,X)] = \int_0^{10-a} (a+x)\lambda e^{-\lambda x} dx$$

  - where $\lambda = 1/5$

$$E[w(a,X)] = \frac{-10\lambda e^{\lambda(a-10)} - e^{\lambda(a-10)} + a\lambda + 1}{\lambda}$$

- Taking derivatives, the maximum is achieved for:

$$a_{opt} = \frac{10\lambda - \log(10\lambda + 1)}{\lambda} = 4.5069\ldots$$

  - and

$$E[w(a_{opt},X)] = a_{opt}$$

Sums of Variables

- Suppose X and Y are independent, what can be said about the distribution of $Z = X+Y$?
- We already know this:

$$P(Z < z) = \int_0^{\infty} P(X < z - y) f_Y(y) dy$$

$$P(Z = z) = \sum_{i=0}^{\infty} P(X = z - i) P(Y = i)$$

  - This is known as known as convolution
- For non-negative variables the density of Z becomes

$$f_Z(z) = \int_0^{z} f_X(z - y) f_Y(y) dy$$

Sums of Variables 2

- Example: convolve two independent exponential variables X and Y:

Erlang(2) distribution

$$f_Z(z) = \int_0^z \lambda e^{-\lambda(z-y)} \lambda e^{-\lambda y} dy = \lambda^2 z e^{-\lambda z}$$

- For the discrete case and non-negative X,Y:

$$P(Z=n) = \sum_{i=0}^{n} P(X=n-i)P(Y=i)$$

- Task:
  - Prove the sum of two Poisson variables with rates λ_1 and λ_2 is another Poisson variable with rate $\lambda_1 + \lambda_2$

Sums of Variables 3

- Convolution is written as *
  - If X has CDF $F(x)$, Y has CDF $G(x)$, then their sum has distribution $F * G$
- It is often needed to establish that the sum $X+Y$ has certain properties, but direct convolution is tedious
  - Such as the sum of binomial variables is binomial
- Or the sum involves many variables
  - Or there is no closed-form solution to the convolution at all and other methods are sought
- An alternative is to use generating functions
  - Similar to Fourier/Laplace transforms
  - Convolution is replaced by multiplication of transform functions, then an inverse transform is applied