Lecture Stat 302
Introduction to Probability - Slides 10

AD

Feb. 2010

Discrete Random Variables

- A discrete r.v. X takes at most a countable number of possible values $\{x_1, x_2, ...\}$ with *p.m.f.*

$$p(x_i) = P(X = x_i)$$

where

$$p(x_i) \geq 0 \text{ and } \sum_{i=1}^{\infty} p(x_i) = 1.$$

- *Expected value/mean*

$$\mu = E(X) = \sum_{i=1}^{\infty} x_i \; p(x_i).$$

- *Variance*

$$Var(X) = E\left((X - \mu)^2\right) = E\left(X^2\right) - \mu^2.$$

Bernoulli and Binomial Distributions

- Assume you have an experiment w.p. success p and w.p. failure $(1-p)$.
- *Bernoulli* r.v.: You set $X=1$ if success and $X=0$ if failure so

$$P(X=1)=p(1)=p,\ P(X=0)=p(0)=1-p.$$

- Assume now you have n independent experiments, each w.p. success p and w.p. failure $(1-p)$.
- *Binomial* r.v.: Set $X=$number of successes among the n experiments, then $X\in\{0,1,2,...,n\}$

$$X=X_1+X_2+\cdots+X_n$$

where X_k is the Bernoulli r.v. associated to experiment i and

$$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}.$$

Example

- *Example*: A survey from Teenage Research Unlimited (Northbrook, Ill.) found that 30% of teenage consumers receive their spending money from part-time jobs. If five teenagers are selected at random, find the probability that at least three of them will have part-time jobs.
- *Answer*: Let X be the number of teenagers having a part-time job among 5 teenagers, then X is Binomial of parameters $n = 5$, $p = 0.3$ so

$$\begin{aligned} P(X \geq 3) &= P(X = 3) + P(X = 4) + P(X = 5) \\ &= 0.132 + 0.028 + 0.002 = 0.162 \end{aligned}$$

- *Example*: What is the probability of obtaining 45 or fewer heads in 100 tosses of a fair coin?
- *Solution*: Let X be the number of heads, then X is binomial of parameters $n = 100$, $p = 0.5$ and

$P(X \leq 45) = \sum_{k=0}^{45} \binom{n}{k} p^k (1-p)^k = 0.184.$

Mean of the Binomial Random Variable

- The mean/expected value of X is given by

$$E(X) = np$$

- **Proof 1**. We have $X = X_1 + X_2 + \cdots + X_n$ where X_i is a Bernoulli r.v. w.p. p so

$$\begin{aligned} E(X) &= E(X_1 + \cdots + X_n) \\ &= E(X_1) + \cdots + E(X_n) = np. \end{aligned}$$

- **Proof 2**. We have

$$\begin{aligned} & E(X) = \sum_{k=0}^{n} k\ P(X = k) = \sum_{k=0}^{n} k \begin{pmatrix} n \\ k \end{pmatrix} p^k (1-p)^{n-k} \\ & = \sum_{k=1}^{n} k\ \frac{n\ (n-1)!}{k\ (k-1)!\ (n-k)!} p\ p^{k-1} (1-p)^{n-k} \\ & = np\ \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!\ (n-k)!}\ p^{k-1} (1-p)^{n-k} \\ & = np\ \sum_{l=0}^{m} \frac{m!}{l!\ (m-l)!}\ p^l (1-p)^{m-l} \quad (m \leftarrow n-1,\ l \leftarrow k-1) \\ & = np \end{aligned}$$

Variance of the Binomial Random Variable

- The variance of the binomial random variable is

$$Var(X) = np(1-p).$$

- **Proof 1**. We have $Var(X) = E(X^2) - E(X)^2 = E(X^2) - (np)^2$ where

$$\begin{aligned} X^2 &= (X_1 + \cdots + X_n)^2 \\ &= \sum_{k=1}^{n} X_k^2 + \sum_{k,l=1;k\neq l}^{n} X_k X_l \end{aligned}$$

so

$$E(X^2) = \sum_{k=1}^{n} E(X_k^2) + \sum_{k,l=1;k\neq l}^{n} E(X_k X_l).$$

We have $E(X_k^2) = 1^2 \times p + 0^2 \times (1-p) = p$. To compute $E(X_k X_l)$ where $k \neq l$, we note that $X_k X_l \in \{0,1\}$ is a r.v. such that

$$\begin{aligned} P(X_k X_l = 1) &= P(X_k = 1) P(X_l = 1) = p^2, \\ P(X_k X_l = 0) &= 1 - P(X_k X_l = 1) = 1 - p^2. \end{aligned}$$

Variance of the Binomial Random Variable

- As a safety check, you have indeed

$$\begin{aligned} P(X_k X_l = 0) &= P(X_k = 0, X_l = 0) + P(X_k = 0, X_l = 1) \\ &\quad + P(X_k = 1, X_l = 0) \\ &= (1-p)^2 + (1-p)p + p(1-p) \\ &= 1 - p^2. \end{aligned}$$

- Hence we have

$$E(X_k X_l) = 1 \times p^2 + 0 \times \left(1 - p^2\right) = p^2$$

thus

$$\begin{aligned} E\left(X^2\right) &= \sum_{k=1}^{n} E\left(X_k^2\right) + \sum_{k=1, l>k}^{n} E(X_k X_l) \\ &= n \times p + n(n-1) \times p^2 \end{aligned}$$

- Finally we have

$$Var(X) = np + n(n-1) \times p^2 - (np)^2$$

Variance of the Binomial Random Variable

- **Proof 2**. We have $Var(X) = E(X^2) - E(X)^2$ where

$$E(X^2) = \sum_{k=0}^{n} k^2 P(X=k) = \sum_{k=1}^{n} k^2 \binom{n}{k} p^k (1-p)^{n-k}$$
$$= \sum_{k=1}^{n} k\ n \binom{n-1}{k-1} p\ p^{k-1} (1-p)^{n-k}$$
$$= \sum_{k=1}^{n} k\ n \binom{n-1}{k-1} p\ p^{k-1} (1-p)^{n-k} = np \sum_{k=1}^{n} k \binom{n-1}{k-1}$$
$$= np \sum_{l=0}^{m} (l+1) \frac{m!}{l!\,(m-l)!} p^l (1-p)^{m-l} = np\ E[Y+1]$$

where Y is Binomial $(n-1, p)$ so $E[Y+1] = (n-1)p + 1$. Finally $Var(X) = np[(n-1)p+1] - (np)^2 = np(1-p)$.

Example: Blood test

- A large number, N, of people are subjected to a blood test. This can be administered in two ways: (1) Each person can be tested separately, in this case N tests are required, (2) the blood samples of k persons can be pooled and analyzed together. If this test is negative, this one test suffices for the k people. If the test is positive, each of the k persons must be tested separately, and in all, $k+1$ tests are required for the k people. Assume that the proba p that a test is positive is the same for all people and that these events are independent.
- (a) Find the probability that the test for a pooled sample of k people will be positive.
- (b) What is the expected value and variance of the number Y of tests necessary under plan (2)? (Assume that N is divisible by k.)

Example: Blood test

- (a) We have

$$\begin{aligned} P\left(\text{pooled sample }+\right) &= P\left(\text{at least one of the } k \text{ people }+\right) \\ &= 1 - P\left(\text{no people }+\right) = 1-\left(1-p\right)^k. \end{aligned}$$

- (b) We have $N = Mk$ so M pooled samples. Let X_i be a Bernoulli variable of parameter $q = 1-\left(1-p\right)^k$ corresponding to success="pooled sample +" then

$$\begin{aligned} Y &= \left(1-X_1\right)+\left(k+1\right)X_1+\left(1-X_2\right)+ \\ &\quad \left(k+1\right)X_2 ... +\left(1-X_M\right)+\left(k+1\right)X_M \\ &= M + k\left(X_1+X_2+\cdots+X_M\right) \end{aligned}$$

where $X = X_1 + X_2 + \cdots + X_M$ is a Binomial of parameters $\left(M, q\right)$. So we have

$$\begin{aligned} E\left(Y\right) &= M + E\left(X\right) = M + kMq, \\ Var\left(Y\right) &= k^2 Var\left(X\right) = k^2 Mq\left(1-q\right). \end{aligned}$$

Poisson Random Variable

- A discrete r.v. X taking values $0, 1, 2, \ldots$ is said to be a Poisson r.v. with parameter λ, $\lambda > 0$, if

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

- This is indeed a probability distribution as

$$\sum_{i=0}^{\infty} p(i) = \sum_{i=0}^{\infty} e^{-\lambda}\frac{\lambda^i}{i!} = e^{-\lambda}\underbrace{\sum_{i=0}^{\infty}\frac{\lambda^i}{i!}}_{=e^{\lambda}} = 1.$$

- This expresses the probability of a number of events occurring in a fixed period of time if these events occur with a *known average rate* λ.
- This is used absolutely everywhere as many natural phenomena are Poisson distributed: the number of soldiers killed by horse-kicks each year in each corps in the Prussian cavalry, the number of phone calls at a call center per minute, the number of murders in London etc.

Example: Misprints in a textbook

- *Example*: There are 50 misprints in a book which has 250 pages and assume these errors follow a Poisson distribution of parameter $\lambda = 50/250 = 0.2$. Find the probability that page 100 has no misprints. Find the probability that page 100 has 2 misprints.
- *Answer*: If we let X be the r.v. denoting the number of misprints on page 100 (or any other page), X is a Poisson r.v. with parameter $\lambda = 0.2$. So we have

$$\begin{aligned}\text{Proba no misprint} &= P(X=0) = e^{-0.2}\frac{(0.2)^0}{0!} = e^{-0.2} = 0.819, \\ \text{Proba 2 misprints} &= P(X=2) = e^{-0.2}\frac{(0.2)^2}{2!} = 0.0164.\end{aligned}$$

Poisson as an approximation to Binomial

- If we consider a binomial r.v. X of parameters (n, p) such that n is large and p is small enough so that np is moderate then the binomial distribution can be well-approximated by the Poisson distribution of parameter $\lambda = np$.
- We have indeed for $\lambda = np \Leftrightarrow p = \frac{\lambda}{n}$

$$
\begin{aligned}
P(X = k) &= \binom{n}{k} p^k (1-p)^{n-k} \\
&= \frac{n!}{k!\ (n-k)!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \\
&= \frac{n(n-1)\cdots(n-k+1)}{n^k} \frac{\lambda^k}{k!} \frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^k} \\
&= 1. \left(1 - \frac{1}{n}\right) \cdots \left(1 - \frac{k-1}{n}\right) \frac{\lambda^k}{k!} \frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^k}
\end{aligned}
$$

Poisson as an approximation to Binomial

- Now recall from your calculus course that

$$\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{n} = e^{-\lambda}, \ \lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{-k} = 1,$$
$$\lim_{n\to\infty}\left(1-\frac{j}{n}\right) = 1$$

so

$$\lim_{n\to\infty} P\left(X=k\right) = e^{-\lambda}\frac{\lambda^{k}}{k!}.$$

- Practically this means that if a large number of trials n are made which have a success probability p such that np is moderate then we can approximate the total number of successes by a Poisson r.v. of parameter $\lambda = np$.

Example: Sampling a population

- A certain disease occurs in 1.2% of the population. 100 people are selected at random from the population. This is a binomial experiment, where the number of trials 100 is large, and the expected number of people with the disease is $\mu = 100 \times 0.12 = 1.2$ is small.
- Assume we want to compute the probability of no people having the disease in a sample of 100. Let X being the number of people having the disease, then X is binomial of parameters $n = 100$ and $p = 1.2/100$ and

$$P(X = 0) = \begin{pmatrix} 100 \\ 0 \end{pmatrix} \left(\frac{1.2}{100}\right)^0 \left(1 - \frac{1.2}{100}\right)^{100} = 0.299016$$

- We can *approximate* X by a Poisson random variable of parameter $\mu = 1.2$. We obtain in this case

$$P(X = 0) = e^{-\mu} \frac{\mu^0}{0!} = .301194$$

Example: Sampling a population

- The probability that exactly 2 people in the sample of 100 have the disease is following the binomial distribution

$$P(X = 2) = \begin{pmatrix} 100 \\ 2 \end{pmatrix} \left(\frac{1.2}{100}\right)^2 \left(1 - \frac{1.2}{100}\right)^{98} = 0.2183$$

whereas the Poisson approximation gives

$$P(X = 2) = e^{-\mu} \frac{\mu^2}{2!} = 0.2169.$$

Example: Random Seeds

- You have a 10 meter by 10 meter plot of land, which is divided into a grid of 100 squares. You scatter 500 seeds on this plot. We assume that each seed falls at random, so that it is equally likely to fall on any of the 100 squares. Consider the square in the upper left hand corner. What is the probability that exactly 4 seeds fall on it? What is the probability that 0 seeds fall on it?
- Think of this as dropping 500 seeds, one after the other, and recording whether the seed falls into the upper left hand square or not. "Success" means falling into the square, and that happens with probability $1/100 = .01$. So, $n = 500$, and $p = 0.1$. The expected number of successes is $np = 5$. By the Poisson approximation, we have

$$P(\text{exactly 4 seeds}) \approx e^{-5}\frac{5^4}{4!} = 0.175.$$

The probability of no seeds is

$$P(\text{no seed}) \approx e^{-5} = 0.0067.$$

Example: Lunch Hour at UBC

- During the noon lunch hour, 47 customers will walk through the door of the post office. Assume that each person arrives at a random time, independent of the other customers. What is the probability that more than one person walks through the door during the first minute?
- To see why this is binomial, think of each of the 47 persons choosing at random a minute during which to arrive. The probability that they choose the first minute is then $1/60$. Thus, we have 47 repetitions of a trial, where the probability of success is $1/60$. The expected number of successes is $np = 47/60 \approx 0.783$.
- To find the probability, we say

$$\begin{aligned} P(\text{more than 1 arrival}) &= 1 - P(\text{no arrival}) \\ &\approx 1 - e^{-0.783} - 0.783e^{-0.783} \\ &\approx 0.185. \end{aligned}$$

Mean of a Poisson Random Variable

- We have $E(X) = \lambda$.
- We have

$$\begin{aligned} E(X) &= \sum_{i=0}^{\infty} i \; P(X = i) = \sum_{i=0}^{\infty} i \; e^{-\lambda} \frac{\lambda^i}{i!} \\ &= e^{-\lambda} \sum_{i=0}^{\infty} \frac{\lambda \; \lambda^{i-1}}{(i-1)!} \\ &= e^{-\lambda} \lambda \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} \quad \text{(change } j \leftarrow i-1\text{)} \\ &= e^{-\lambda} \lambda e^{\lambda} = \lambda. \end{aligned}$$

- Note that this is in agreement with our approximation of Binomial by Poisson. A Binomial has mean np and we approximate it by a Poisson of parameter $\lambda = np$ which is also the mean of the Poisson distribution.

Variance of a Poisson Random Variable

- We have $Var(X) = E(X^2) - E(X)^2 = \lambda$.
- Proof: We have

$$\begin{aligned} E\left(X^2\right) &= \sum_{i=0}^{\infty} i^2 \; e^{-\lambda} \frac{\lambda^i}{i!} = \sum_{i=1}^{\infty} i^2 \; e^{-\lambda} \frac{\lambda^i}{i!} \\ &= e^{-\lambda} \sum_{i=1}^{\infty} i \; \frac{\lambda \; \lambda^{i-1}}{(i-1)!} = \lambda \sum_{j=0}^{\infty} (j+1) \; e^{-\lambda} \frac{\lambda^j}{j!} \\ &= \lambda E(Y+1) \end{aligned}$$

  where Y is a Poisson random variable of parameter λ, hence $E(Y+1) = \lambda + 1$.
- We can now conclude

$$Var(X) = \lambda(\lambda+1) - \lambda^2 = \lambda.$$

- Note that for Poisson random variable, we have $E(X) = Var(X) = \lambda$. For binomial we have $E(X) = np$ and $Var(X) = np(1-p)$ and so $Var(X) \approx E(X)$ only if $p << 1$.

Example: Waiting at the bus stop

- *Example*: Assume that 6 buses per hour stop at your bus stop. If the buses were arriving exactly every 10 minutes, then your average waiting time would be 5 minutes; i.e. you wait between 0 and 10 minutes. What is the probability that you are going to wait at least 5 minutes without seeing any bus if the buses follow a Poisson distribution? What is the proba to wait at least 10 minutes? What is the proba of seeing two buses in 10 minutes?
- *Answer*: If we let X_1 be the number of buses arriving in 5 minutes, it is a Poisson r.v. with parameter 0.5 (average rate 6 per hour). So we have

$$P(X_1 = 0) = e^{-0.5} = 0.60.$$

If we let X_2 be the number of buses arriving in 10 minutes, it is a Poisson r.v. with parameter 1. So we have

$$P(X_2 = 0) = e^{-1} = 0.368, \; P(X_2 = 2) = e^{-1}\frac{1^2}{2!} = 0.184.$$

Example: Murders in London

- Each year the London Metropolitan Police record around 160 murders, and this has been stable for the last 5 years. Each of these murders is an individual crime that cannot be predicted. It may appear strange, but this very randomness means that the overall pattern of murders is in some ways quite predictable.
- Assuming the number of murders are Poisson distributed, compute the probability of having no murder during a day, 3 or more murders in one day, a week without any murder.

Example: Murders in London

- We have an average rate per day of $\lambda = 160/365 \approx 0.44$.
- *Peaceful day* (no murder). The proba that of having no murder is $P(X=0) = e^{-\lambda} = 0.64$.
- *Bloody day* (3 or more murders). The proba is $P(X \geq 3) = 1 - e^{-\lambda} - \lambda e^{-\lambda} - \frac{\lambda^2}{2} e^{-\lambda} = 0.0103 \approx \frac{3.75}{365}$
- *Peaceful week* (no murder over a week). We have a Poisson distribution $\lambda' = 160/365 \times 7$

$$P(X=0) = e^{-\lambda'} \approx 0.0465 \approx \frac{2.4}{52}.$$