Math 502AB - Lecture 7

Dr. Jamshidian

September 13, 2017

1 Lecture - Part 1

1.1 Poisson Approximation To Binomial

Let $X \sim binomial(n, p)$. Then:

$$\begin{aligned}\frac{f(k|n,p)}{f(k-1|n,p)} &= \frac{(n-k+1)p}{k(1-p)} \\ &= \frac{np-kp+p}{k-kp}\end{aligned}$$

- The way the poisson approximation to the binomial holds is when $n \to \infty$, $p \to 0$, and $np \to \lambda$

$$\frac{f(k|n,p)}{f(k-1|n,p)} = \frac{\lambda}{k} \quad (*)$$

- For large n and small p: $np \equiv \lambda$, $p \equiv \frac{\lambda}{n}$

$$f(0|n,p) = (1-p)^n = \left(1 - \frac{\lambda}{n}\right)^n \to e^{-\lambda}$$

Recall: $Y \sim poisson(\lambda)$ means:

$$f_Y(k|\lambda) = \frac{e^{-\lambda}\lambda^k}{k!}$$

So we now have the convergence to $e^{-\lambda} = f_Y(0|\lambda)$. Similarly, we have:

$$f(1|n,p) \cong \lambda f(0|n,p) \to \lambda e^{-\lambda} = f_Y(1|\lambda)$$

$$f(2|n,p) \cong \frac{1}{2}\lambda f(1|n,p) \to \frac{1}{2}\lambda^2 e^{-\lambda} = f_Y(2|\lambda)$$

And, by induction, you can show that, in general:

$$f(k|n,p) \to \frac{1}{k!}\lambda^k e^{-\lambda} = f_Y(k|\lambda)$$

1.2 Deriving the Poisson *pmf*

$$f_Y(k|\lambda) = \frac{e^{-\lambda}\lambda^k}{k!}$$

Consider events that occur at random points in time, and consider the number of events in a given time interval. In particular, consider an interval of length t and divide this interval into n equal sub-intervals.

(Insert cool number line photo here)

Let $h = \frac{t}{n}$. We would like to show that the number of events in the interval $[0, t]$ has a *poisson* distribution.

Assumptions:

1. The probability that exactly one event occurs in a given interval of length h is equal to $\lambda h + o(h) \quad \left[\frac{o(h)}{h} \to 0\right]$. In other words, this is linear in h. The $o(h)$ term goes to zero faster than h
2. For small h, the probability that two or more events occur in an interval of length h is $o(h)$
3. Occurrences of events in disjoint sub-intervals are independent

Under these assumptions, we will show that the number of events in the interval $[0, t]$ has a *poisson* distribution with parameter λt.

- Let $N(t)$ be the number of events in the interval $[0, t]$. We now want to find:
$$P[N(t) = k] = ?$$
- Let $E = k$ of the intervals contain exactly one event, and the remaining $n - k$ intervals contain zero events.
- The above implies E^c is at least one interval containing two or more events.

$$P[N(t) = k] = P\{[N(t) = k] \cap E\} + P\{[N(t) = k] \cap E^c\}$$

Letting the first intersection equal A and the second intersetion equal B, we can see that $B \subset E^c$ implies:

$$\begin{aligned}
P(B) &\leq P(E^c) \\
&= P\left[\bigcup_{i=1}^{n} \text{Interval } i \text{ contains two or more events}\right] \\
&\leq \sum_{i=1}^{n} P\left(\text{Interval } i \text{ contains two or more events}\right) \\
&= \sum_{i=1}^{n} o(h) = \sum_{i=1}^{n} o\left(\frac{t}{n}\right) \to 0 \quad \text{as } n \to \infty
\end{aligned}$$

Thus, $P(B) \to 0$ as $n \to \infty$.

We know that $A = E$ since $E \subset \{N(t) = k\}$. $P(E)$ is the probability that k of the intervals contain exactly one event and the remaining $n-k$ intervals contain 0 events.

$$P(\text{Exactly one event occurs in an interval}) = \lambda h + o(h)$$

$$\begin{aligned} P(\text{0 Events in an interval}) &= 1 - P(\text{1 in an interval}) - P(\text{2+ in an interval}) \\ &= 1 - (\lambda h + o(h)) - o(h) = 1 - \lambda h + o(h) \end{aligned}$$

There are $\binom{n}{k}$ possibility for intervals which contain exactly 1 event:

$$\binom{n}{k}(\lambda h + o(h))^k (1 - \lambda h + o(h))^{n-k} \sim binomial(n, \lambda h + o(h))$$

In this formula we have $p = \lambda h + o(h)$.

$$\begin{aligned} \Rightarrow np &= n\lambda h + no(h) - n\lambda \frac{t}{n} + no\left(\frac{t}{n}\right) \\ p &= \lambda h + o(h) = \lambda\left(\frac{t}{n}\right) + o\left(\frac{t}{n}\right) \to 0 \end{aligned}$$

Thus the claim that the *poisson* distribution models "counts" effectively holds *only when* the assumptions in **1.2** hold.

2 Lecture - Part 2: Discrete Distributions

2.1 Negative Binomial

$$E(Y) = r\left(\frac{1-p}{p}\right), \quad Var(Y) = \frac{r(1-p)}{p^2}$$

We will show that as $r \to \infty$, $p \to 1$, and $r(1-p) \to \lambda$, then:

$$NBinom(r, p) \to Poisson(\lambda)$$

2.2 Geometric Distribution

The *geometric distribution* is a special case of the *negative binomial* where $r = 1$, making X "the number of trials until your first success".

$$f_X(x|p) = p(1-p)^{x-1} \quad x = 1, 2, ...$$

$$E(X) = \frac{1}{p}, \quad Var(x) = \frac{1-p}{p^2}$$

2.2.1 Memoryless property of the *geometric distribution*

For integers $s > t$, we have:

$$P(X > s | X > t) = P(X > s - t)$$

Proof:

$$\begin{aligned}
P(X > s | X > t) &= \frac{P(X > s \cap X > t)}{P(X > t)} \\
&= \frac{P(X > s)}{P(X > t)} \\
&= \frac{(1-p)^s}{(1-p)^t} = (1-p)^{s-t} \\
&= P(X > s - t)
\end{aligned}$$

Example:

Let X be the number of samples required to reach a smoker, with the probability of success p. If $s = 10$ and $t = 6$ then:

$$P(X > 10 | X > 6) = P(X > 4)$$

3 Lecture - Part 2: Continuous Distributions (Ch 3.3)

3.1 Uniform Distribution

We say that $X \sim unif[a, b]$ if:

$$f_X(x|a, b) = \begin{cases} \frac{1}{b-a} & a \leq X \leq b \\ 0 & \text{otherwise} \end{cases}$$

$$E(X) = \frac{b+a}{2}, \quad Var(x) = \frac{(b-a)^2}{12}$$

3.2 Gamma Distribution

$$f(x|\alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta} \quad x, \alpha, \beta > 0$$

3.2.1 Gamma-Poisson Relationship

Let α be an integer, $X \sim gamma(\alpha, \beta)$, and $Y \sim poisson\left(\frac{x}{\beta}\right)$ for any given x. Then

$$P(X \leq x) = P(Y \geq \alpha)$$

Proof (by induction)

$\alpha = \mathbf{1}$:

$$
\begin{aligned}
P(X \leq x) &= \int_0^x \frac{1}{\Gamma(1)\beta^2} t^{1-1} e^{-t/\beta} dt \\
&= \int_0^x \frac{1}{\beta} e^{-t/\beta} = 1 - e^{-x/\beta} \\
&= P(Y \geq 1)
\end{aligned}
$$

Suppose the statement holds for $\alpha = n - 1$. That is $P(X \leq x) = P(Y \geq n - 1)$. We need to show that the statement holds for $\alpha = n$

$$P(X \leq x) = P(Y \geq n)$$

Suppose that $\alpha = n$

$$P(X \leq x) = \int_0^x \frac{1}{\Gamma(n)\beta^\alpha} t^{n-1} e^{-t/\beta}$$

And... Using integration by parts (let $u = t^{n-1}$ and $dv = e^{-t/\beta}dt$), we can prove it.

3.3 Exponential distribution

A special case of $gamma(\alpha, \beta)$ is the $exponential(\beta)$ distribution.

$$X \sim gamma(\alpha = 1, \beta) \iff X \sim exponential(\beta)$$

The exponential random variable hs memoryless property as well:

$$P(X > s | X > t) = P(X > s - t)$$

3.4 Weibull Distribution

If $X \sim exp(\beta)$ then $Y = X^{1/\gamma}$, with (γ ¿ 0), is said to have a $weibull(\gamma, \beta)$ distribution:

$$F_Y(y|\gamma, \beta) = \frac{\gamma}{\beta} y^{\gamma - 1} e^{-y^2/\beta} \quad y, \gamma, \beta > 0$$

3.5 Normal (Gaussian) Distribution

If $X \sim N(\mu, \sigma^2)$, then $Z = \frac{x-\mu}{\sigma} \sim N(0, 1)$. Thus it is important to know that:

$$M_Z(t) = e^{t^2/2}$$

3.6 Beta Distribution

We say that $X \sim Beta(\alpha, \beta)$ if:

$$f(x|\alpha, \beta) = \frac{1}{\mathcal{B}(\alpha, \beta)} x^{\alpha-1}(1-x)^{\beta-1} \quad 0 < x < 1, a, \beta > 0$$

Where $\mathcal{B}(\alpha, \beta)$ is the beta function as defined by:

$$\mathcal{B}(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

If $\alpha = \beta$, you get a symmetric distribution about $x = \frac{1}{2}$. With $E(X) = \frac{1}{2}$ and $Var(X) = \frac{1}{8\alpha+4}$

3.7 Cauchy Distribution

If $X \sim Cauchy(\theta)$, then:

$$f(x|\theta) = \frac{1}{\pi} \cdot \frac{1}{1 + (x-\theta)^2} \quad -\infty < x < \infty$$

$E(X)$ does not exist.