MATH 425 10/24/2011

Note Title 10/24/2011

The expectation and variance of the hypergeometric random variable

N balls m white n chosen without replacement

$X = \#$ chosen white

$$P\{X=i\} = \frac{\binom{m}{i}\binom{N-m}{n-i}}{\binom{N}{n}}$$

To compute expectation and variance, compute a general (recursive) formula for the moment. (Example 8g p. 162)

$$\binom{m}{i} = \frac{m}{i}\binom{m-1}{i-1} \qquad \binom{N}{n} = \frac{N}{n}\binom{N-1}{n-1}$$

reduce n, m, N, i by 1

$k > 0$

$$E(X^k) = \sum_{i=0}^{n} i^k \, P\{X = i\} = \sum_{i=0}^{n} i^k \binom{m}{i}\binom{N-m}{n-i} \Big/ \binom{N}{n} =$$

$j = i - 1$

$$= \sum_{i=1}^{n} i^{k-1} \, m \binom{m-1}{i-1} \binom{(N-1)-(m-1)}{(n-1)-(i-1)} \Big/ \frac{N}{n}\binom{N-1}{n-1} =$$

$$= \frac{mn}{N} \sum_{j=0}^{n-1} (j+1)^{k-1} \binom{m-1}{j} \binom{(N-1)-(m-1)}{(n-1)-j} \Big/ \binom{N-1}{n-1}$$

$$= \frac{mn}{N} E\left((Y+1)^{k-1}\right) = \frac{mn}{N} \sum_{l=0}^{k-1} \binom{k-1}{l} E(Y^l)$$

Let Y be the hypergeometric with parameters $n-1, m-1, N-1$

$$\boxed{E(X^k) = \frac{mn}{N} \sum_{l=0}^{k-1} \binom{k-1}{l} E(Y^l)}$$

$k = 1$

$$E(X) = \frac{mn}{N}$$

(note: same as for the binomial distribution with $p = \frac{m}{N}$, n Bernoulli trials)

$$E(X^2) = \frac{mn}{N}\left(E(Y) + 1\right) = \frac{mn}{N}\left(\frac{(m-1)(n-1)}{N-1} + 1\right)$$

$$\mathrm{var}(X) = E(X^2) - E(X)^2 = \frac{mn}{N}\left(\frac{(m-1)(n-1)}{N-1} + 1 - \frac{mn}{N}\right)$$

$\boxed{\text{put } p = \frac{n}{N}}$

$$= np\left(\frac{\cancel{np} - p}{\cancel{(n-1)p}} - (n-1)\frac{1-p}{N-1} + 1 - \cancel{np}\right)$$

$$\frac{n-1}{N-1} = \frac{Np-1}{N-1} = p - \frac{1-p}{N-1}$$

$$= \underbrace{np(1-p)}\left(1 - \frac{(n-1)}{N-1}\right)$$

variance of the binomial distribution

Oh, by the way, $E(X+Y) = E(X) + E(Y)$, ...

CONTINUOUS RANDOM VARIABLES

(Recall: discrete random variable meant: countably many values)

An (absolutely) continuous random variable X on a sample space S is one for which there exist a real function $f : \mathbb{R} \to [0, \infty)$ such that for every measurable subset B of $\mathbb{R}$,

$$P\{X \in B\} = \int_B f(x)\, dx.$$

(Real analysis explains: measurable, Lebesgue integral 597.)

It is enough to assume

$$P\{X \le a\} = \int_{-\infty}^{a} f(x)\, dx. \quad \text{for every } a \in \mathbb{R}.$$

$f(x)$ is a continuous analogue of the probability mass function, and is called the <u>probability density function</u>.

What is expectation and variance for continuous random variables?

$$E(X) = \int_{-\infty}^{\infty} x f(x)\, dx$$

<u>Note</u>:

$$\int_{-\infty}^{\infty} f(x)\, dx = 1.$$

Higher moments:

$$E(X^k) = \int_{-\infty}^{\infty} x^k f(x)\,dx$$

$$\text{var}(X) = E(X^2) - E(X)^2.$$

Example: Let X be a continuous random variable with probability density function

$$f(x) = \begin{cases} 1 & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

(uniform distribution).

Compute $E(X)$, $\text{var}(X)$.

Solution:

$$E(X) = \int_0^1 x\,dx = \frac{x^2}{2}\Big|_0^1 = \frac{1}{2}$$

density zero outside of the bounds

$$E(X^2) = \int_0^1 x^2\,dx = \frac{x^3}{3}\Big|_0^1 = \frac{1}{3}$$

$$E(X^2) - E(X)^2 = \frac{1}{3} - \frac{1}{4} = \underline{\underline{\frac{1}{12}}}.$$

Example 2b p.190: Using the random variable X from the previous example, compute $E(e^X)$.

Solution: $E(e^X) = \int_0^1 e^x \, dx = e^x \Big|_0^1 = e - 1.$

$$\left(E(g(X)) = \int_{-\infty}^{\infty} g(x) f(x) \, dx \right) \leftarrow \text{Proposition 2.1 p. 191}$$

when X has probability density function $f(x)$.

The general uniformly distributed random variable X:

a random choice of a number in an interval $[a, b]$
(unbiased)

Probability density: $f(x) = \begin{cases} \frac{1}{b-a} & x \in [a,b] \\ 0 & x \in [a,b] \end{cases}$

Expected value: $\int_a^b \frac{1}{b-a} x \, dx = \frac{1}{b-a} \frac{x^2}{2} \Big|_a^b =$

$$= \frac{b^2-a^2}{b-a} \cdot \frac{1}{2} = \frac{1}{2}(a+b) = E(X)$$

Variance: $E(X^2) = \int_a^b \frac{1}{b-a} x^2 dx = \frac{1}{b-a} x^3 \Big|_a^b =$

$$= \frac{1}{b-a} (b^3-a^3) \cdot \frac{1}{3} = (a^2+ab+b^2) \cdot \frac{1}{3}$$

$$\mathrm{var}(X) = E(X^2) - E(X)^2 = (a^2 + ab + b^2)\frac{1}{3} - (a^2 + 2ab + b^2)\frac{1}{4}$$

$$= \frac{(a-b)^2}{12}. \qquad \left(\frac{1}{12}(a^2+b^2) - \frac{1}{6}ab\right)$$

HW: (1) 5.1 p.224

(2) 5.14 p.225

$$\text{Distribution} = \int_{-\infty}^{a} f(x)\,dx$$

$$= P\{X \le a\}$$

Quiz Wednesday (discrete variables)