## 3.1 The Binomial Distribution

Def'n: A Bernoulli trial is a trial with only two possible outcomes, usually termed as a "success" and a "failure". The prob. of success is $p$ and the prob. of failure is $1 - p$.
Consider a random experiment consisting of $n$ Bernoulli trials such that

1. The trials are independent.
2. Each trial results in only 2 possible outcomes: "success" and "failure".
3. The probability of success in each trial ($p$) remains constant.

Let an r.v. $X$ be the number of successful trials while $0 < p < 1$ and $n = 1, 2, \ldots$
Then, an $X$ with these parameters has a binomial distribution and its *pmf* is

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x} \qquad x = 0, 1, \ldots, n$$

Also,

$$\mu = E(X) = np \qquad \text{and} \qquad \sigma^2 = V(X) = np(1-p)$$

Ex3.1) Recall Oilers ticket phone line example from Ch. 1. Let $X$ be the number of calls succeeding in making a sale. Thus, $X \sim B(n, p) = B(10, 0.3)$

a) What is the probability that exactly 5 calls of 10 result in a sale?

$$P(X = 5) = \binom{n}{x} p^x (1-p)^{n-x} =$$

b) What is the probability that at least 3 calls result in a sale being made?

c) What is the probability that more than 3 but no more than 5 calls result in a sale?

d) What are the mean and standard deviation of the number of sales?

### 3.2.1 Geometric Distribution

Def'n: Consider still a series of Bernoulli trials (independent trials with constant prob. $p$ of success); here, however, there is no need to define $n$. Let the r.v. $X$ be the number of trials until the 1st success with parameter $0 < p < 1$. Then, $X$ has a geometric dist'n with

$$f(x) = (1-p)^{x-1} p \qquad x = 1, 2, \ldots$$

Also,

$$\mu = E(X) = \frac{1}{p} \qquad \text{and} \qquad \sigma^2 = V(X) = \frac{(1-p)}{p^2}$$

Ex3.2) The probability of a random engineer consuming alcohol during a weekend is 0.95. Assume all engineers are independent.

a) At a party, what is the probability of the third engineer you meet being the first passed out on the floor? Let $X$ denote the number of engineers found until the lush is found.

b) Find the mean and standard deviation of $X$.

$E(X) =$

$\sigma =$

Note that the independence of trials means the count of the # of trials can be started at any trial without changing the probability distribution of $X$. For example, if you've found 20 engineers and you start over, the probability of finding the first convivial engineer *after* engineer #20 is the same probability as finding the initial lush. Since $p$ remains constant, the geometric distribution exhibits the *lack of memory property*.

$$P(X < t + \Delta t \mid X > t) = \frac{P(t < X < t + \Delta t)}{P(X > t)} = \ldots = P(X < \Delta t)$$

| | Binomial | Geometric |
|---|---|---|
| # of trials | Constant ($n$) | Variable ($X$) |
| # of successes | Variable ($X$) | 1 |

### 3.2.2 Negative Binomial Distribution

Def'n: Consider again a series of Bernoulli trials (independent trials with constant probability $p$ of success) and there remains no need to define $n$. Let the r.v. $X$ be the number of trials until $r$ successes occur with parameters $0 < p < 1$ and $r = 1, 2, 3, \ldots$. Then, $X$ has a negative binomial distribution with

$$f(x) = \binom{x-1}{r-1} (1-p)^{x-r} p^r \qquad x = r, r+1, r+2, \ldots$$

Note that the special case of $r = 1$ is a geometric random variable.

Also,

$$\mu = E(X) = \frac{r}{p} \qquad \text{and} \qquad \sigma^2 = V(X) = \frac{r(1-p)}{p^2}$$

| | Binomial | Negative binomial |
|---|---|---|
| # of trials | Constant ($n$) | Variable ($X$) |
| # of successes | Variable ($X$) | Constant ($r$) |

Ex3.3) Assume the Oilers winning percentage of 0.409 (from last season) is $p$ and each game is an independent event. Let $X$ denote the number of games played until the 20th win so that $r = 20$.

a) What is the probability that it takes 55 games to get 20 wins? 65 games?

$$P(X = 55) = \binom{x-1}{r-1}(1-p)^{x-r}\,p^r =$$

$$P(X = 65) = \binom{x-1}{r-1}(1-p)^{x-r}\,p^r =$$

b) What is the probability that it takes at least 23 games to get 20 wins?

c) Find the mean and standard deviation of $X$.

$E(X) =$

$\sigma =$

3.4 Poisson Distribution

Def'n: Given an interval of real numbers, assume events occur at random throughout the interval. The random variable $X$ that equals the # of successes in the interval has a Poisson distribution with parameter $\lambda > 0$ and a *pmf* that is

$$f(x) = \frac{e^{-\lambda}\lambda^x}{x!} \qquad x = 0, 1, 2, \ldots$$

Also, if $X \sim Poisson(\lambda)$,

$$\mu = E(X) = \lambda \qquad \text{and} \qquad \sigma^2 = V(X) = \lambda$$

Ex3.4) In the last NHL season, suppose the number of goals in a game follows a Poisson distribution of 5.51 goals per game. Let $X$ be the number of goals in a game.

a) What is the probability of exactly 8 goals in a game? ($\lambda = 5.51$ goals/game)

$$P(X=8) = \frac{e^{-\lambda}\lambda^x}{x!} =$$

b) What is the probability of 1 or fewer goals in a game?

c) What is the probability of exactly 2 goals in 10 minutes?

d) What is the probability of 30 goals in 4 games?

e) What are the mean and standard deviation for $X \sim Poisson(\lambda)$?

If the interval can be partitioned into subintervals small enough such that

1. the prob. of more than one event in a subinterval is approx. zero,
2. the prob. of one event in a subinterval is the same for all subintervals and proportional to the length of the subinterval, and
3. the event in each subinterval is independent of other subintervals,

the random experiment is called a Poisson process.

(diagram drawn in class)

In other words, the Poisson process is a binomial experiment with *infinite n* trials. Recall that for the binomial distribution, $E(X) = np$ (while $E(X) = \lambda$ for Poisson). Then,

$$\lim_{n\to\infty} B(n,p) = \lim_{n\to\infty}\binom{n}{x}p^x(1-p)^{n-x} = \lim_{n\to\infty}\binom{n}{x}\left(\frac{\lambda}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^{n-x} = \frac{e^{-\lambda}\lambda^x}{x!}$$

For example, Ex3.4 "could" be a Poisson process if an appropriate subinterval is found.

| | Binomial | Poisson |
|---|---|---|
| # of trials | Constant ($n$) | Infinite |
| # of successes | Variable ($X$) | Variable ($X$) |
| Prob. of success | Constant ($p$) | Constant ($p = \lambda/n$) |