Chapter 5 Special Probability Distributions

§ 5.1 Introduction

In this chapter we introduce several common pdf's associated with discrete and continuous r.v.'s.

§ 5.2 The Discrete Uniform Distribution

If a discrete r.v. takes on k values with equal probability we say it has a discrete uniform distribution.

<u>Defn</u> A r.v. X has a discrete uniform distribution iff its pdf is given by

$$f(x) = \frac{1}{k}, \quad x = x_1, x_2, \ldots, x_k$$

where $x_1, x_2, \ldots, x_k$ are all distinct.

In this case we have

$$\mu = E(X) = \sum_{i=1}^{k} x_i \cdot \frac{1}{k}$$

$$\sigma^2 = V(X) = \sum_{i=1}^{k} (x_i - \mu)^2 \cdot \frac{1}{k}.$$

An important special case is when $x_i = i$, $1 \le i \le k$.

Here
$$\mu = \sum_{i=1}^{k} \frac{i}{k} = \frac{1}{k} \cdot \frac{k(k+1)}{2} = \frac{k+1}{2}$$

$$\mu_2' = E(X)^2 = \sum_{i=1}^{k} \frac{i^2}{k} = \frac{1}{k} \cdot \frac{k(k+1)(2k+1)}{6} = \frac{(k+1)(2k+1)}{6}$$

So
$$\sigma^2 = \mu_2' - \mu^2 = \frac{(k+1)(2k+1)}{6} - \frac{(k+1)^2}{4} = \frac{k^2-1}{12}$$

§ 5.3 Bernoulli Distribution

Suppose I flip a biased coin once where the chance is θ I get a H and $1-\theta$ I get a tail (here $0 < \theta < 1$). This is an example of a Bernoulli random variable.

Defn. A r.v. X has a Bernoulli (θ) distribution for some $\theta \in (0,1)$ iff its pdf satisfies

$$f(x;\theta) = \begin{cases} 1-\theta, & x=0 \\ \theta, & x=1 \\ 0 & \text{otherwise.} \end{cases}$$

Can also write this as

$$f(x;\theta) = \theta^{x}(1-\theta)^{1-x}, \quad x=0,1.$$

Thus $f(0; \theta) = 1-\theta$

$f(1; \theta) = \theta$

and we can calculate

$$\mu = E(X) = 0 \cdot (1-\theta) + 1 \cdot \theta = \theta$$

$$\sigma^2 = V(X) = \mu_2' - \mu^2$$

$$= 0 \cdot (1-\theta) + 1^2 \cdot \theta - \theta^2$$

$$= \theta(1-\theta).$$

This will be useful to us later when we deal with the next type of pdf, the binomial distribution.

§ 5.4 Binomial Distribution

Suppose now I flip the same biased coin of the last section n times. What is the prob. that in the n flips I get x heads (where $0 \leq x \leq n$)?

Well, the chance of getting x heads and $n-x$ tails <u>in a particular order</u> is

$$\theta^x (1-\theta)^{n-x}$$

and there are $\binom{n}{x}$ suitable orderings.

Thus, the prob. I get exactly x heads in n flips is

$$\binom{n}{x} \theta^x (1-\theta)^{n-x}.$$

Def^n A r.v. has a binomial (n, θ) distribution $(n \in \mathbb{N},\ 0 < \theta < 1)$ iff its pdf is given by

$$b(x; n, \theta) = \binom{n}{x} \theta^x (1-\theta)^{n-x}, \quad x = 0, 1, 2, \ldots, n.$$

Note that if $n = 1$, $\binom{n}{x} = 1$ for $x = 0, 1$ and

$$b(x; 1, \theta) = \theta^x (1-\theta)^{1-x}, \quad x = 0, 1$$

and we recover the Bernoulli distribution (as we would expect - why?). Thus the binomial distribution generalizes the Bernoulli distribution.

N.b. the n flips (trials / experiments) must be independent if we are to get a binomial (n, θ) r.v. from n Bernoulli (θ) r.v.'s.

Ex. The prob. of getting 5H & 7T in 12 flips of a balanced coin is

$$b(5; 12, \tfrac{1}{2}) = \binom{12}{5} \cdot \left(\tfrac{1}{2}\right)^5 \left(1-\tfrac{1}{2}\right)^{12-5}$$

$$= \binom{12}{5} \cdot \left(\frac{1}{2}\right)^{12}$$

(Here $n = 12$, $\theta = \frac{1}{2}$, $x = 5$).

Ex. Find the prob that 7 out of 10 people recover from a tropical disease if we assume indep. and the prob that any given person recovers is 0.8.

Here $n = 10$, $\theta = .8$, $x = 7$ and so the answer is

$$b(7; 10, 0.8) = \binom{10}{7}(.8)^7(.2)^3 \approx .2$$

Back to coin flips for a while ---

If I flip the biased coin n times, then, since we are measuring the <u>total</u> number of successes (heads) in the n trials, if $X_1, X_2, \ldots, X_n$ are n indep Bernoulli (θ) r.v.'s, one for each flip, then

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

is binomial (n, θ).

Here the r.v.'s $X_1, \ldots, X_n$ are said to be <u>independent identically distributed</u> (iid), an important term we will meet later.

In general, r.v.'s $Y_1, \ldots, Y_n$ are iid if they are indep & have the same pdf.

This representation of a binomial (n, θ) r.v as the sum of n iid Bernoulli (θ) r.v's is useful for calculating the mean and variance of this pdf.

Indeed, recalling from the last section that the Bernoulli (θ) distrib. has mean θ & variance $\theta(1-\theta)$, we get

$$\mu = E(X) = E(X_1 + \cdots + X_n)$$

$$= \sum_{i=1}^{n} E(X_i) \qquad \text{(Thm 4.3, p. 133)}$$

$$= \sum_{i=1}^{n} \theta \qquad \text{by iid}$$

$$= n\theta$$

$$\sigma^2 = V(X) = V(X_1 + \cdots + X_n)$$

$$= \sum_{i=1}^{n} V(X_i) \quad \text{by Cor 4.3, p 154 as } X_1, \ldots, X_n \text{ are indeps}$$

$$= \sum_{i=1}^{n} \theta(1-\theta) \quad \text{by iid.}$$

$$= n\theta(1-\theta).$$

We have proved

<u>Theorem 5.2</u> The mean and variance of the binomial (n, θ) distrib. are

$$\mu = n\theta, \qquad \sigma^2 = n\theta(1-\theta).$$

n.b. This can also be calculated directly from

$$b(x; n, \theta) = \binom{n}{x} \theta^x (1-\theta)^{n-x}$$

- see the book.

Another way to obtain the mean & variance is via the MGF.

<u>Thm 5.4</u> The MGF of the binomial (n, θ) distribution is given by

$$M_X(t) = [1 + \theta(e^t - 1)]^n.$$

Pf.
$$M_X(t) = E(e^{Xt})$$
$$= \sum_x e^{xt} f(x)$$
$$= \sum_{x=0}^{n} e^{xt} \binom{n}{x} \theta^x (1-\theta)^{n-x}$$
$$= \sum_{x=0}^{n} \binom{n}{x} (\theta e^t)^x (1-\theta)^{n-x}$$
$$= [\theta e^t + (1-\theta)]^n \quad \text{by the Binomial thm.}$$
$$= [1 + \theta(e^t - 1)]^n. \quad \blacksquare$$

If we diff $M_X(t)$ twice wrt t, then

$$M_X'(t) = n\theta e^t \left[1+\theta(e^t-1)\right]^{n-1}$$

$$M_X''(t) = n\theta e^t \left[1+\theta(e^t-1)\right]^{n-1}$$
$$+ n(n-1)\theta^2 e^{2t}\left[1+\theta(e^t-1)\right]^{n-2}$$
$$= n\theta e^t(1-\theta+n\theta e^t)\left[1+\theta(e^t-1)\right]^{n-2}$$

And if we subst $t=0$, we get by Thm 4.9 p. 145

$$\mu = \frac{d}{dt} M_X(t)\Big|_{t=0} = n\theta$$

$$\mu_2' = \frac{d^2}{dt^2} M_X(t)\Big|_{t=0} = n\theta(1-\theta+n\theta)$$

So that

$$\sigma^2 = \mu_2' - \mu^2 = n\theta(1-\theta+n\theta) - n^2\theta^2$$
$$= n\theta(1-\theta)$$

as before.

§ 5.5 The Negative Binomial and Geometric Distributions

Back to the repeated Bernoulli trials (coin flips).

Suppose we want to know the prob. that it will take x trials until we have k successes.

Well, the last trial must be a success and for a Bernoulli (θ) r.v. the prob. of this success is θ.

Of the remaining $x-1$ trials which went before, exactly $k-1$ must be successes and as we saw with the binomial distr, the prob. of this is

$$\binom{x-1}{k-1} \theta^{k-1} (1-\theta)^{(x-1)-(k-1)}$$

Since the trials are <u>independent</u>, to find the prob that the kth success occurs on the xth trial, we multiply these probs to get

$$\theta \binom{x-1}{k-1} \theta^{k-1} (1-\theta)^{(x-1)-(k-1)}$$

$$= \binom{x-1}{k-1} \theta^{k} (1-\theta)^{x-k}$$

<u>Defn</u> A r.v. X has a negative binomial(k, θ) distribution iff its pdf is given by

$$b^*(x; k, \theta) = \binom{x-1}{k-1} \theta^{k} (1-\theta)^{x-k}$$

$$x = k, k+1, \ldots$$

$$(0 < \theta < 1).$$

Ex. If the prob. is .4 that a child exposed to a certain contagious disease will catch it, what is the prob. that the 10^{th} child exposed will be the 3^{rd} to catch the disease?

Soln. Here $x = 10$, $k = 3$, $\theta = .4$ and the answer is

$$b^*(10; 3, .4) = \binom{9}{2}(.4)^3(.6)^7$$

$$= .0645$$

We have the following relationship between the negative binomial and the binomial.

Thm 5.5

$$b^*(x; k, \theta) = \frac{k}{x} \cdot b(k; x, \theta).$$

One can show the following

Thm 5.6 The mean & variance of the negative binomial distn. are

$$\mu = \frac{k}{\theta} \quad \text{and} \quad \sigma^2 = \frac{k}{\theta}\left(\frac{1}{\theta} - 1\right).$$

An important special case of the negative binomial is the geometric distribution

Here $k=1$ and we want the prob. that the first success occurs on the xth trial.

Defn A r.v. X has a geometric distribution iff its pdf is given by

$$g(x;\theta) = \theta(1-\theta)^{x-1}, \quad x = 1, 2, 3, \ldots$$

Note that the mean & variance of the geom distn. are

$$\mu = \frac{1}{\theta}, \quad \sigma^2 = \frac{1}{\theta}\left(\frac{1}{\theta} - 1\right).$$

by Thm 5.6 with $k = 1$.

<u>Ex 5.5</u> If the prob. is .75 that an applicant for a driver's licence will pass the road test on any given try, what is the prob the applicant will pass the test on the 4th attempt?

<u>Soln</u> Here $x = 4$, $\theta = .75$ and the answer is

$$g(4; .75) = .75\,(1 - .75)^{4-1}$$

$$= .75\,(.25)^3$$

$$= .0117.$$

§ 5.7 The Poisson Distribution

Useful for calculating the number of successes when we have many trials and the chance of success in each trial is low (and the same).

Another way of saying this is that the Poisson distribution is useful for calculating the number of occurrences of rare events (e.g. road accidents).

<u>Def^n</u>. A r.v. X has a <u>Poisson (λ) distribution</u> iff its pdf is given by

$$p(x;\lambda) = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x = 0, 1, 2, \ldots$$

Note that $p(x;\lambda) \geqslant 0$ everywhere and

$$\sum_x p(x;\lambda) = \sum_{x=0}^{\infty} \frac{\lambda^x e^{-\lambda}}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!}$$

$$= e^{-\lambda} \cdot e^{\lambda} = \underline{1}$$

and so we do have a genuine pdf.

<u>In practice</u> if we have n trials with prob θ of success in each, then if n is large and θ small and we set $\lambda = n\theta$, then the prob. of x successes is exactly

$$\binom{n}{x} \theta^x (1-\theta)^{n-x}$$

is well approximated by

$$\frac{\lambda^x e^{-\lambda}}{x!} \qquad (\lambda = n\theta).$$

Ex If 2% of the books at a certain bindery have defective bindings, find the (approx) prob. that 5 out of 400 books bound by this bindery have defective bindings.

Soln. Here $x = 5$, $\lambda = n\theta = 400 \times .02 = 8$

and the answer is

$$p(5;8) = \frac{8^5 \cdot e^{-8}}{5!} \approx .093.$$

Ex. Records show that the prob. is .00005 that a car will have a flat tyre crossing the bridge to Newport. Use the Poisson distrib to approx. the binomial probs. that among 10,000 cars crossing this bridge

a) exactly one will have a flat tyre;

b) at most two will have a flat tyre.

Soln. a) Here $x = 2$, $\lambda = 10,000(.00005) = .5$

and the answer is

$$p(2; .5) = \frac{(.5)^2 e^{-.5}}{2!} \approx .0758$$

b) In terms of events

{at most 2 cars have a flat tyre}

= {no car has a flat tyre}

$\cup$

{exactly one car has a flat tyre}

$\cup$

{exactly two cars have a flat tyre}

and this union is (pairwise) disjoint.

We already know the prob of the third event. For the first two, $x = 0, 1$ and the resp. probs. using $p(x; .5)$ are approx $.6065$ and $.3033$ resp.

The total prob that at most 2 cars have a flat tyre is then approx

$$.6065 + .3033 + .0758 = .9856$$

In order to calculate the mean and variance of the Poisson distrib., we will use the MGF.

<u>Theorem 5.9</u> The MGF of the Poisson (λ) distribution is

$$M_X(t) = e^{\lambda(e^t - 1)}.$$

Pf.

$$M_X(t) = E(e^{tX})$$

$$= \sum_x e^{xt} \cdot f(x)$$

$$= \sum_{x=0}^{\infty} e^{xt} \cdot \frac{\lambda^x e^{-\lambda}}{x!}$$

$$= e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!}$$

$$= e^{-\lambda} \cdot e^{\lambda e^t}$$

$$= e^{\lambda(e^t-1)}. \quad \blacksquare.$$

If we now diff twice wrt t, we get

$$M_X'(t) = \lambda e^t e^{\lambda(e^t-1)}$$

$$M_X''(t) = \lambda e^t e^{\lambda(e^t-1)} + \lambda^2 e^{2t} e^{\lambda(e^t-1)}$$

Setting $t=0$ and using Thm 4.9 P.145 as before gives

$$\mu = \mu_1' = M_X'(0) = \lambda$$

$$\mu_2' = M_X''(0) = \lambda + \lambda^2$$

$$\sigma^2 = \mu_2' - \mu^2 = \lambda + \lambda^2 - \lambda^2$$
$$= \lambda.$$

We have proved

<u>Thm 5.8</u> The mean and variance of the Poisson (λ) distrib are given by

$$\mu = \lambda, \quad \sigma^2 = \lambda.$$

Since the expectation of a Poisson (λ) distrib is λ, we can use it to model the number of (rare) events in a given time interval, given that we already know the average number of events in that much time.

Ex. The average no. of trucks arriving on any one day at a truck depot is 12. What is the prob that on a given day fewer than 9 trucks arrive at the depot?

Soln. Here $\lambda = 12$ and the answer is

$$P(X<9) = \sum_{x=0}^{8} p(x;12) \approx .1550.$$

Suppose we observe on average 2 accidents per hour on a certain stretch of I-95.

If we watched for 2 hours (or 2 separate 1 hour intervals), we would expect an average of 4 accidents

If we watched for 3 hours, we would expect an average of 6 accidents, and so on.

In general, if the average number of successes per unit time/area is, α, then on a time interval of length t or a region of area α, we would expect αt successes.

On this interval, we would have a Poisson (αt) r.v. with distrib.

$$p(x;\alpha t) = \frac{e^{-\alpha t}(\alpha t)^x}{x!}, \quad x = 0,1,2,\ldots$$

Ex. A certain kind of sheet metal has on average 5 defects per $10 ft^2$. If we assume a Poisson distrib, what is the prob. that a $15 ft^2$ sheet will have at least 6 defects?

Soln. Let X = # defects in a $15 ft^2$ sheet.

Since the (unit) area is $10 ft^2$, we have that X is Poisson (λ) with

$$\lambda = 5(1.5) = 7.5.$$

Then

$$P(X \geq 6) = 1 - P(X \leq 5)$$

$$= 1 - P(X=0) - P(X=1) - P(X=2) - P(X=3) - P(X=4) - P(X=5)$$

$$= 1 - \frac{e^{-7.5}(7.5)^0}{0!} - \frac{e^{-7.5}(7.5)^1}{1!} - \frac{e^{-7.5}(7.5)^2}{2!} - \frac{e^{-7.5}(7.5)^3}{3!} - \frac{e^{-7.5}(7.5)^4}{4!} - \frac{e^{-7.5}(7.5)^5}{5!}$$

$$\approx 0.7586.$$

§ 5.8 Multinomial Distribution

We consider a generalization of the repeated coin flips where there are several possible outcomes.

To be specific suppose we have n independent trials with k mutually exclusive outcomes where resp. probs. are $\theta_1, \theta_2, \ldots, \theta_k$ so that

$$\sum_{i=1}^{k} \theta_i = 1.$$

Suppose we want

x_1 outcomes of the first kind

x_2 - - second - -

$\vdots$

x_k - - kth - -

where of course $x_1 + x_2 + \cdots + x_k = n$.

The chance of getting these outcomes <u>in a particular order</u> is

$$\theta_1^{x_1} \theta_2^{x_2} \cdots \theta_k^{x_k}$$

while the number of possible (favourable) orderings is given by the multinomial coefficient

$$\binom{n}{x_1, x_2, \ldots, x_k} = \frac{n!}{x_1! \, x_2! \cdots x_k!}$$

<u>Defn</u> The r.v.'s $X_1, \ldots, X_k$ have a multinomial $(n, \theta_1, \ldots, \theta_k)$ distribution iff their joint pdf is given by

$$f(x_1, \ldots, x_k; n, \theta_1, \ldots, \theta_k) = \binom{n}{x_1, \ldots, x_k} \theta_1^{x_1} \cdots \theta_k^{x_k}$$

for $x_i = 0, \ldots, n$ with $\sum_{i=1}^{n} x_i = n$, $\sum_{i=1}^{k} \theta_i = 1$.

Ex. A certain city has 3 TV stations. On Sat evening Channel 12 has 50% of the viewing audience, Ch 10 has 30% and Ch 3 has 20%.

Find the prob that among 8 TV viewers in the city, 5 will be watching Ch 12, 2 will be watching Ch 10 and 1 will be watching Ch 3.

Soln. Here $n = 8$, $x_1 = 5$, $x_2 = 2$, $x_3 = 1$

$\theta_1 = .5$, $\theta_2 = .3$, $\theta_3 = .2$ and the answer is

$$f(5, 2, 1; 8, .5, .3, .2)$$

$$= \frac{8!}{5!\,2!\,1!} (.5)^5 (.3)^2 (.2)$$

$$= .0945$$