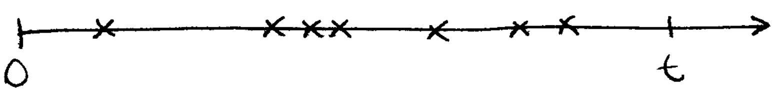
14.381 HANDOUT #2
FALL 2004

TA: JOSÉ TESSADA
(tessada@mit.edu)
OH: Th., 4–6 pm.

1. POISSON POSTULATES



probability that a certain number of events occurr between 0 and t.

NOTICE THAT THE TIME IS CONTINUOUS!

(Thm. 3.8.1) For each $t \geq 0$, let N_t be an integer-valued random variable with the following properties. (Think of N_t as denoting the number of arrivals in the time period from time 0 to time t.)

(i) $N_0 = 0$ (start with no arrivals)

(ii) $s < t \Rightarrow N_s$ and $N_t - N_s$ are independent (arrivals in disjoint time periods are indep.)

(iii) N_s and $N_{t+s} - N_t$ are identically distributed (number of arrivals depend only on period length)

(iv) $\lim_{t \to 0} \dfrac{P(N_t = 1)}{t} = \lambda$ (arrival prob. is proportional to period length, if length is small)

(v) $\lim_{t \to 0} \dfrac{P(N_t > 1)}{t} = 0$ (no simultaneous arrivals)

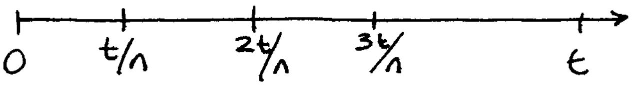
If (i)–(v) hold, then for any integer n,

$$P(N_t = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!},$$

that is, $N_t \sim \text{Poisson}(\lambda t)$

Notes: • also has applications in spatial distributions.

Heuristic proof:



probability of success: $\frac{\lambda t}{n}$ [by (iv)]

$$\cancel{\approx P\left(N_{\frac{it}{n}} - N_{\frac{(i-1)t}{n}}\right)}$$

$$P\left(N_{\frac{it}{n}} - N_{\frac{(i-1)t}{n}} > 1\right) \approx 0$$

$$P\left(N_t = k\right) \approx \binom{n}{k} \left(\frac{\lambda t}{n}\right)^k \left(1 - \frac{\lambda t}{n}\right)^{n-k}$$

If we increase the number of subdivisions we obtain a finer grid and we can obtain the poisson pdf as the limit when $n \to \infty$

$$P(N_t = k) \approx \frac{n(n-1)\cdot \ldots \cdot (n-k+1)}{k!\, n^k} (\lambda t)^k \left(1 - \frac{\lambda t}{n}\right)^{n-k}$$

$$\frac{n(n-1)\cdot \ldots \cdot (n-k+1)}{n^k} \longrightarrow 1 \quad \text{as} \quad n \to \infty$$

$$\left(1 - \frac{\lambda t}{n}\right)^n \longrightarrow e^{-\lambda t} \quad \text{as} \quad n \to \infty$$

$$\left(1 - \frac{\lambda t}{n}\right)^{-k} \longrightarrow 1 \quad \text{as} \quad n \to \infty$$

so $$\binom{n}{k}\left(\frac{\lambda t}{n}\right)^k \left(1 - \frac{\lambda t}{n}\right)^{n-k} \longrightarrow \frac{1}{k!} (\lambda t)^k e^{-\lambda t}$$

2. LOCATION AND SCALE FAMILIES

We "construct" families of distributions starting from a single pdf $f(x)$, the standard pdf for the family.

Most famous example: Normal distribution with μ and σ^2.

Thm. 3.5.1: Let $f(x)$ be any pdf and let μ and $\sigma > 0$ be any given constants. Then the function

$$g(x|\mu,\sigma) = \frac{1}{\sigma} f\left(\frac{x-\mu}{\sigma}\right)$$

is a pdf.

Proof: we check that $g(x) > 0$ for every value of μ and σ, and that integrates to 1.

$\Rightarrow$ · as $f(x)$ is a pdf, then for $\sigma > 0$, $g(x) \geq 0$.

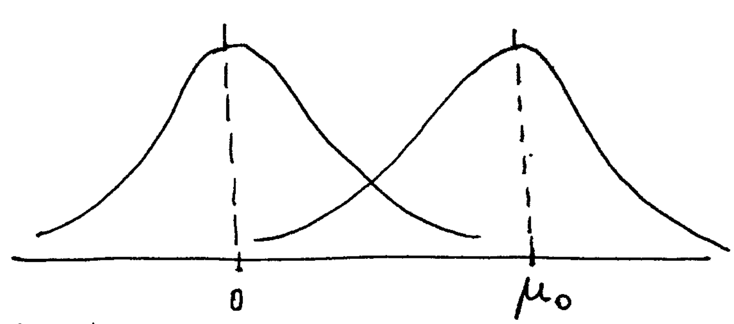
·
$$\int_{-\infty}^{\infty} \frac{1}{\sigma} f\left(\frac{x-\mu}{\sigma}\right) dx = \int_{-\infty}^{\infty} f(y)\, dy \qquad \text{change of variable: } y = \frac{x-\mu}{\sigma}$$

$$= 1$$

■

Definition 3.5.2: Let $f(x)$ be any pdf. Then the family of pdfs $f(x-\mu)$, indexed by the parameter μ, $-\infty < \mu < \infty$, is called the location family with standard pdf $f(x)$ and μ is called the location parameter for the family.

Graphically:



it is true that X is a r.v. with pdf $f(x)$ and $\tilde{X}$ is a r.v. with pdf $f(\tilde{x})$

then

$$P(-a \le X \le b \mid 0) = P(\mu_0 - a \le \tilde{X} \le \mu_0 + b \mid \mu_0)$$

where $a, b > 0$ are constants.
So we just move the distribution from one place to other, but "the way it looks" doesn't change!

Definition 3.5.4: Let $f(x)$ be any pdf. Then for any $\sigma > 0$, the family of pdfs $(1/\sigma) f(x/\sigma)$, indexed by the parameter σ, is called the scale family with standard pdf $f(x)$ and σ is called the scale parameter of the family.

example: normal if $\mu = 0$
gamma if α is fixed and β is the scale
(this includes the exponential for the case $\alpha = 1$).

Definition 3.5.5: Let $f(x)$ be any pdf. Then for any μ, $-\infty < \mu < \infty$, and any $\sigma > 0$, the family of pdfs $(1/\sigma) f((x-\mu)/\sigma)$, indexed by the parameter (μ, σ), is called the location-scale family with standard pdf $f(x)$; μ is called the location parameter and σ is called the scale parameter.

$\sigma > 1 \longrightarrow$ "stretch"
$\sigma < 1 \longrightarrow$ "contract"

example: normal dist.

Thm. 3.5.6: Let $f(\cdot)$ be any pdf. Let μ be any real number, and let σ be any positive real number. Then X is a random variable with pdf $(1/\sigma) f((x-\mu)/\sigma)$ iff $\exists$ a r.v. Z with pdf $f(z)$ and $X = \sigma Z + \mu$.

(see also Thm. 3.5.7).

3. EXPONENTIAL FAMILIES

Definition: A (parametric) family $\{f(\cdot|\theta) : \theta \in \Theta\}$ of pmfs/pdfs is an exponential family if

$$f(x|\theta) = h(x)\, c(\theta) \exp\left[\sum_{i=1}^{k} w_i(\theta)\, t_i(x)\right]$$

$h : \mathbb{R} \to \mathbb{R}_+$ $\quad$ $w_i : \Theta \to \mathbb{R}$

$c : \Theta \to \mathbb{R}_{++}$ $\quad$ $t_i : \mathbb{R} \to \mathbb{R}$

This is going to be useful in statistics

Exercise 3.28 parts (d) and (e)

Show that each of the following families is an exponential fam.:

(d) Poisson family

$$f(x|\theta) = \exp(-\lambda) \cdot \frac{\lambda^x}{x!} \quad \text{if } x \in \mathbb{N}_0$$

$$h(x) = \frac{1}{x!} \cdot \mathbb{1}(x \in \{0,1,2,\dots\})$$

$c(\theta) = e^{-\theta}$ $\quad$ $w_1(\theta) = \log\theta$ $\quad$ $t_1(x) = x$

(e) negative binomial family with r known, $0 < p < 1$

$$P(X = x|p) = \binom{x-1}{r-1} p^r (1-p)^{x-r} \quad x = r, r+1, \dots$$

$$h(x) = \binom{x-1}{r-1} \mathbb{1}(\{r, r+1, \dots\} \ni x)$$

$$c(p) = \left(\frac{p}{1-p}\right)^r \qquad w_1(p) = \log(1-p) \qquad t_1(x) = x$$

Note: this is not true if r is unknown

4. CHEBYCHEV'S INEQUALITY

Thm. 3.6.1: Let X be a random variable and let $g(x)$ be a nonnegative function. Then, for any $r > 0$

$$P(g(X) \geq r) \leq \frac{E[g(X)]}{r}$$

Example 3.6.2: $g(x)=(x-\mu)^2/\sigma^2$, where $\mu = E[X]$
$\sigma^2 = Var[X]$

write $r=t^2$, then

$$P\left(\frac{(X-\mu)^2}{\sigma^2} \geq t^2\right) \leq \frac{1}{t^2} E\left[\frac{(X-\mu)^2}{\sigma^2}\right] = \frac{1}{t^2}$$

notice that σ^2 is a constant and $E[(X-\mu)^2]=\sigma^2$

Doing some algebra we can express the previous inequality in the following way:

$$P(|X-\mu| \geq t\sigma) \leq 1/t^2$$

or $$P(|X-\mu| < t\sigma) \geq 1-\frac{1}{t^2}$$

for $t=2$
$$P(|X-\mu| \geq 2\sigma) \leq 1/2^2 = 0.25$$

for $t=1$?

5. MULTIPLE RANDOM VARIABLES

$$f_{X,Y}(x,y) = \frac{1}{x} \qquad 0 \leq y \leq x \leq 1$$

are X and Y independent?

Let's find $f_X(x)$ and $f_{Y|X}(y|x)$

$$f_X(x) = \int_0^x f(x,y)\,dy = \int_0^x \frac{1}{x}\,dy = \left(y/x\right)\Big|_0^x = 1$$

$$f_X(x) = 1, \qquad 0 \le x \le 1$$

$$f_{Y|X}(y|x) = 1/x \qquad 0 \le y \le x \le 1$$

So,

- X uniformly distributed on $[0,1]$
- conditional on $X = x$, Y is uniform on $[0, x]$

→ Constructing the joint pdf ...

Suppose we have two random variables X, Y.
$X \sim N(\mu_X, \sigma_X^2)$, $Y \sim N(\mu_Y, \sigma_Y^2)$.
We can think of $f_X(x)$ and $f_Y(y)$ as being the marginal distributions, can we infer something about the joint pdf?

If we assume independence, yes!
If that's the case, then

$$f_{X,Y}(x,y) = f_X(x)\, f_Y(y) \qquad (x,y) \in \mathbb{R}^2$$

This is very usual in practice.

$f_{X,Y}(x,y)$ built in this way is just a particular case of the bivariate normal distribution.