PROBABILITY DISTRIBUTIONS [CONTINUED]

CONTINUOUS RANDOM VARIABLES

Very often we need to work with random variables that may [in principle] vary continuously over some range of values.

E.g. human lifespan [0 to < 200 years]
masses of asteroids
length of rats' tails
time till a nucleus decays [0 to ∞]

If our measuring device [metre rule?] has a resolution Δx [1 cm divisions?] we could define a DISCRETE probability distribution

$$P_i = P(x_i < x < x_i + \Delta x)$$

E.g. for rats tails $P_5 = P(5 < x < 6\text{ cm}) \neq 0.$

As the resolution Δx is improved [1 mm divisions on rule?], the fraction of the population in range x to $x + \Delta x$ will decrease, reaching zero in limit $\Delta x \to 0$:

$P(x = 6\text{ cm}) = 0$ [no rat has a tail of exactly 6 cm!]

[CONTINUOUS RAND. VARS., CONTD]

To avoid the problem of vanishing probabilities, we introduce the PROBABILITY DENSITY FUNCTION [or DISTRIBUTION FUNCTION] $f(x)$ via

$$P(x_i < X < x_i + \Delta x) = f(x_i)\,\Delta x \quad [\Delta x \to 0]$$

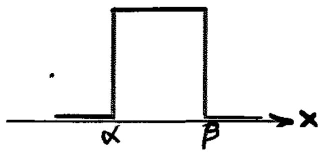
so that

$$P(a < x < b) = \lim_{\Delta x \to 0} \sum_{\substack{i \\ x_i \in (a,b)}} f(x_i)\,\Delta x = \int_a^b f(x)\,dx$$

NOTE: $f(x) \geq 0$ and $\int_{-\infty}^{\infty} f(x)\,dx = 1.$

Unlike PROBABILITIES [which are dimensionless], DISTRIBUTION FUNCTIONS have dimensions inverse to those of the variable x [since $f(x)\,\Delta x$ must be dimensionless]; e.g., $[\text{length}]^{-1}$, $[\text{mass}]^{-1}$, etc.

EXAMPLE 1: The UNIFORM DISTRIBUTION



$f(x) = 0$ for $x < \alpha$ or $x > \beta$

$f(x) = \dfrac{1}{\beta - \alpha}$ otherwise.

EXPECTATION VALUE

Extend the definition from Lecture 4:

$$E(x) \equiv \langle x \rangle = \lim_{\Delta x \to 0} \sum_i x_i \underbrace{f(x_i)\Delta x}_{\text{"}p_i\text{"}} = \int_{-\infty}^{\infty} x\, f(x)\, dx$$

EXAMPLE 2: the UNIFORM DISTRIBUTION —

$$\langle x \rangle = \int_{-\infty}^{\infty} x\, f(x)\, dx = \int_\alpha^\beta x \cdot \frac{1}{\beta-\alpha}\, dx = \left[\frac{\frac{1}{2}x^2}{\beta-\alpha}\right]_\alpha^\beta = \frac{\frac{1}{2}(\beta^2-\alpha^2)}{\beta-\alpha} = \frac{1}{2}(\alpha+\beta)$$

VARIANCE & STANDARD DEVIATION

Similarly,

$$\sigma^2 \equiv \mathrm{var}(x) = \langle x^2 \rangle - \langle x \rangle^2$$

EXAMPLE 3: the UNIFORM DISTRIBUTION again —

$$\langle x^2 \rangle = \int_{-\infty}^{\infty} x^2 f(x)\, dx = \frac{1}{\beta-\alpha}\int_\alpha^\beta x^2\, dx = \frac{\frac{1}{3}(\beta^3-\alpha^3)}{(\beta-\alpha)} = \frac{1}{3}(\alpha^2+\alpha\beta+\beta^2).$$

$$\Rightarrow \sigma^2 = \langle x^2 \rangle - \langle x \rangle^2 = \frac{\alpha^2+\alpha\beta+\beta^2}{3} - \left(\frac{\alpha+\beta}{2}\right)^2 = \frac{1}{12}(\beta-\alpha)^2.$$

← [Hard work! Can you see a quicker way?]

CUMULATIVE PROBABILITY

... defined by

$$C(x) \equiv P(Y < x) = \int_{-\infty}^{x} f(y)\, dy.$$

↖ the random variable! (Y)

Note that $C(-\infty) = 0$, $C(+\infty) = 1$ and that

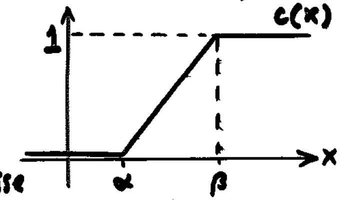
$$\frac{dC}{dx} = f(x) \geqslant 0 \quad [C \text{ never decreases}]$$

EXAMPLE 4: for the uniform distribution,

$C(x) = 0$ for $x < \alpha$

$C(x) = 1$ for $x > \beta$

$$C(x) = \int_\alpha^x \frac{dx}{\beta-\alpha} = \frac{x-\alpha}{\beta-\alpha} \text{ otherwise}$$



It's often useful to define a "SURVIVAL PROBABILITY", $P(x) \equiv 1 - C(x)$

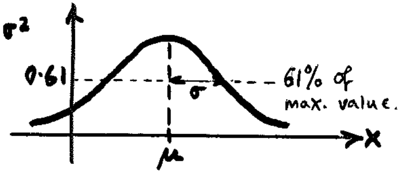
$$= \int_x^\infty f(y)\, dy,$$

e.g., the probability that a nucleus survives to time x [or that a rat's tail is longer than x?]

Clearly, $\frac{dP}{dx} = -\frac{dC}{dx} = -f(x) \leqslant 0$: the probability of survival decreases with time...

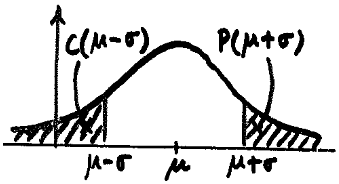
THE GAUSSIAN [or NORMAL] DISTRIBUTION

$$f(x) = \frac{e^{-(x-\mu)^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}}$$



Its importance [e.g. in the theory of "errors"] arises from the CENTRAL LIMIT THEOREM of probability theory: the distribution function for the SUM of many independent random variables is a Gaussian [almost] regardless of the distribution for any one term in the sum. [The experimental "error" often arises from many independent sources.]

The "cumulative" functions $C(x)$ & $P(x)$ are usually looked up in tables (!), and are useful in assessing the "significance" of an expt. result [i.e., its "degree of improbability"]



$C(\mu-\sigma) = P(\mu+\sigma) \doteq 16\%$

$C(\mu+2\sigma) = P(\mu-2\sigma)$
$\doteq 2{\cdot}3\%$

etc.

INTEGRALS OF GAUSSIANS [DIGRESSION]

In cases where the range is infinite [so not $P(x)$, $C(x)$ or similar...] the key result is

$$\int_{-\infty}^{\infty} e^{-ax^2}\,dx = \left(\frac{\pi}{a}\right)^{1/2}.$$

All others of the form $\int_{-\infty}^{\infty} x^{2n} e^{-ax^2}\,dx$ can be obtained by differentiation w.r.t. a:

E.g.
$$\int_{-\infty}^{\infty} x^2 e^{-ax^2}\,dx = \int_{-\infty}^{\infty}\left[-\frac{de^{-ax^2}}{da}\right]dx$$
$$= -\frac{d}{da}\left[\int_{-\infty}^{\infty} e^{-ax^2}\,dx\right]$$
$$= -\frac{d}{da}\left(\frac{\pi}{a}\right)^{1/2} = \frac{\pi^{1/2}}{2\,a^{3/2}} \qquad (E)$$

EXAMPLE 5: variance of Gaussian distribution

$f(x) = \exp[-(x-\mu)^2/2\sigma^2]/[\text{whatever}]$

where $[\text{whatever}] = \int_{-\infty}^{\infty} \exp[-(x-\mu)^2/2\sigma^2]\,dx$.

Set $a \equiv 1/(2\sigma^2)$

$\Rightarrow [\text{whatever}] = (\pi/a)^{1/2}$

$\Rightarrow \text{var}(x) = \langle (x-\mu)^2 \rangle$

$$= \frac{1}{(\pi/a)^{1/2}} \int_{-\infty}^{\infty} (x-\mu)^2 e^{-a(x-\mu)^2}\,dx$$

[CONTINUED]

[EXAMPLE 5, CONTINUED]

By making a change of variable $u = x - \mu$ we find that the required integral is the same as the one evaluated in (E):

$$\mathrm{Var}(x) = \frac{1}{(\pi/a)^{1/2}} \int_{-\infty}^{\infty} u^2 e^{-au^2}\, du$$

$$= \frac{1}{(\pi/a)^{1/2}} \left\{ \frac{\pi^{1/2}}{2a^{3/2}} \right\} = \frac{1}{2a}$$

$$= \sigma^2,$$

i.e., the parameter σ is the <u>standard deviation</u> of the Gaussian distribution.