pfsl8.tex

Lecture 8. 24.10.2013 (half-hour: Problems)

We form the *probability generating function (PGF)*

$$P(s), \text{ or } P_X(s), := E[s^X] = \sum_{n=0}^{\infty} s^n P(X = n) = \sum_{n=0}^{\infty} p_n s^n.$$

This is a power series in s, and since $\sum p_n = 1$, it converges for $s = 1$. So the radius of convergence R is at least 1.

If $R > 1$, $P(s)$ is analytic (= holomorphic) at $s = 1$, so we may differentiate termwise:

$$P'(s) = \sum_{n=1}^{\infty} n s^{n-1} p_n; \qquad P''(s) = \sum_{n=2}^{\infty} n(n-1) s^{n-2} p_n.$$

Taking $s = 1$:

$$P'(1) = \sum n p_n = \sum n P(X = n) = E[X];$$

$$P''(1) = \sum n(n-1) p_n = \sum n(n-1) P(X = n) = E[X(X-1)],$$

etc. (the right-hand sides are called the *factorial moments* of X; they determine the moments, and vice versa). Thus

$$E[X] = P'(1)];$$

$$var(X) = E[X^2] - (E[X])^2 = E[X(X-1)] + E[X] - (E[X])^2$$
$$= P''(1) + P'(1) - [P'(1)]^2.$$

This gives the mean and variance in terms of the first two derivatives of the PGF, in the case $R = 1$. We quote that these formulae still hold even if $R = 1$. This depends on Abel's Continuity Theorem from Analysis; we omit this.

Convolution.

Just as with MGFs and CFs: the PGF of an independent sum is the product of the PGFs.

Random sums.

If we have a random sum – a sum $X_1 + \ldots + X_N$ of a random number N of iid random variables X_i, where the X_i have PGF $P(s)$ and N is independent of the X_i with PGF $Q(s)$ – then $X_1 + \ldots + X_N$ has PGF $Q(P(s))$, the functional composition of P and Q. This result is very useful in the study of *branching processes*, which model the growth of biological populations (or chain reactions in Physics, Chemistry etc.); see Problems 10 Q3.

III. CONVERGENCE and LIMIT THEOREMS

1. Modes of convergence.

In Analysis, we deal with convergence and limits all the time, but in Probability Theory we have to modify our requirements.

Example: Coin tossing. Consider repeated (independent) tosses of a fair coin (outcomes iid Bernoulli $B(\frac{1}{2})$). What can we say about the long-run behaviour of the observed frequency to heads to date? The man/woman in the street will say, "tends to a half – Law of Averages". There is much good sense in this, and we will prove a theorem that says just this, but *subject to a qualification*, that turns out to be inevitable.

The coin can fall tails (frequency of heads 0; pr $\frac{1}{2}$). So it can fall tails 10 times (frequency of heads 0; pr 2^{-10}); 100 times (frequency 0; pr 2^{-100}), etc. Such highly exceptional behaviour is certainly very unusual (highly unlikely – and we can say exactly how unlikely). In the limit, we would expect the probability of this or any other aberrant behaviour to tend to 0, and it does. The fact remains that the limit of 0 is 0, and *not* the $\frac{1}{2}$ occurring in the Law of Averages.

Because of such examples, the best we can hope for is the following.

Definition. We say that random variables X_n, $n = 1, 2, \ldots$, *converge to* X *almost surely (a.s.)*, or *with probability 1 (wp1)*, if

$$P(X_n \to X \ (n \to \infty)) = 1,$$

and write

$$X_n \to X \qquad a.s.$$

This is one of our two *strong* modes of convergence. For the other:

Definition. For $p \geq 1$, X_n *converges to* X *in pth mean*, or *in* L_p, if

$$E[|X_n - X|^p] \to 0 \qquad (n \to \infty)$$

(L for Lebesgue, p for pth power). The two most important cases for us are $p = 1$ – convergence *in mean*, or *in* L_1, and $p = 2$ – convergence *in mean square*, or *in* L_2.