Module 10.4: Binomial Distribution

MCIT Online - CIT592 - Professor Val Tannen

Lecture Notes

Binomial random variables

An r.v. $B : \Omega \to \mathbb{R}$ is called **binomial** with parameters $n \in \mathbb{N}$ and $p \in [0,1]$ when $\mathrm{Val}(B) = [0..n]$ and $\forall k \in [0..n] \quad \Pr[B = k] = \binom{n}{k} p^k (1-p)^{n-k}$.

How does such an r.v. arise? For example, perform n IID Bernoulli trials with probability of success p and let B be the r.v. that returns the number of successes observed. Clearly $\mathrm{Val}(B) = [0..n]$. Then:

We have seen before the probability space on which B is defined: the outcomes are the 2^n sequences of length n of $\mathtt{S}$'s (for "success") and $\mathtt{F}$'s (for "failure"). An outcome with k $\mathtt{S}$'s has probability $p^k(1-p)^{n-k}$.

There are $\binom{n}{k}$ outcomes with k $\mathtt{S}$'s. Therefore the probability of the event "k successes observed" is $\binom{n}{k} p^k (1-p)^{n-k}$.

The distribution of B is also called **binomial** with parameters n and p.

Examples of binomial

Example. We considered the r.v.'s X_H and X_T that return the number of heads, and respectively tails, shown when we flip a fair coin twice.

Both X_H and X_T are binomial r.v.'s with $n = 2$ and $p = 1/2$.

Example. We throw k balls into m bins. The r.v. that returns the number of balls that end up in Bin 1 is a binomial r.v. with $n = k$ and $p = 1/m$.

Example. We roll a fair die k times. The r.v. that returns the number of twos and threes shown is a binomial r.v. with $n = k$ and $p = 1/3$.

Expectation for binomial

Since expectation (and variance) are completely determined by the distribution it suffices to compute them for B be the (binomial) r.v. that returns the number of successes observed in n IID **Bernoulli** trials, each with probability of success p.

Let S_i be the event "the i'th trial resulted in S". We have $\Pr[S_i] = p$.

Let J_i be the **indicator** random variable of the event S_i.
$\mathrm{E}[J_i] = \Pr[J_i = 1] = \Pr[S_i] = p$

Now, the binomial r.v. B can be **decomposed** into a sum of indicators:
$B = J_1 + \cdots + J_n$.

By linearity of expectation $\mathrm{E}[B] = \mathrm{E}[J_1] + \cdots + \mathrm{E}[J_n]$.

In conclusion $\mathrm{E}[B] = p + \cdots + p = np$.

Variance for binomial I

Yet again, let B be the (binomial) r.v. that returns the number of successes observed in n IID **Bernoulli** trials, each with probability of success p.

On the previous slide we decomposed $B = J_1 + \cdots + J_n$ where J_i is the **indicator** random variable of the event $S_i =$"the i'th trial resulted in S".

In general, variance does **not** distribute over sums. However, we have

Proposition. If r.v.'s $X_1, \ldots, X_n$ are **pairwise independent** then

$$\mathrm{Var}[X_1 + \cdots + X_n] \;=\; \mathrm{Var}[X_1] + \cdots + \mathrm{Var}[X_n]$$

The **proof** (and a more general formulation) is in the segment "Correlated random variables".

Variance for binomial II

$B = J_1 + \cdots + J_n$ where (recall) J_i is the indicator random variable of the event S_i = "the i'th trial resulted in S".

The events S_i are mutually independent and therefore the indicator r.v.'s J_i are also mutually independent. We can apply the proposition:

$$\mathrm{Var}[J_1 + \cdots + J_n] = \mathrm{Var}[J_1] + \cdots + \mathrm{Var}[J_n]$$

We have calculated earlier the variance of Bernoulli r.v.'s with parameter p.

Therefore, $\mathrm{Var}[J_i] = \Pr[J_i = 1](1 - \Pr[J_i = 1]) = p(1-p)$.

In conclusion, $\mathrm{Var}[B] = p(1-p) + \cdots + p(1-p) = np(1-p)$.