Probability Basics: Handout 3

Math/Stat 394: Probability I

Wellner; 1/21/2000

SUMMARY SHEET: THE BERNOULLI PROCESS

$X_i \equiv$ (success or failure indicator for i-th trial) $\sim Bernoulli(p)$.
$P(X_i = k) = p^k(1-p)^{1-k}, k = 0, 1.$

$T_n \equiv$ $X_1 + \cdots + X_n =$ (the total number of successes in the first n trials)
$\sim Binomial(n, p)$

$Y_i \equiv$ (the i-th interarrival time) $\sim Geometric(p)$
$P(Y_i = k) = q^{k-1}p$ for $k = 1, 2, \cdots$ with
$E(Y_i) = 1/p$ and $Var(Y) = q/p^2$.

$W_r \equiv$ $Y_1 + \cdots + Y_r$ (the waiting time until the r-th success)
$\sim NegativeBinomial(r, p)$;
$P(W_r = k) = \binom{k-1}{r-1} q^{k-r}p^r$, for $k = r, r+1, \cdots$
with $E(W_r) = r/p$ and $Var(W_r) = rq/p^2$.
We must have exactly $r-1$ failures in the first $k-1$ trials, and we must have a sucess on the k-trial. There are $\binom{r-1}{k-1}$ such sequences of 0's and 1's, and each one has probability $q^{k-r}p^r$.

- Key fact: $[W_r > n] = [T_n < r]$.
- Binomial facts:

  (a) The basic Bernoulli rv X has mean $\mu_X = p$ and variance $\sigma_X^2 = p(1-p) \equiv pq$.

  (b) The rv T_n has mean $\mu_n = np$ and variance $\sigma_n^2 = npq$.

  (c) The rv

$$Z_n \equiv \frac{T_n - \mu_n}{\sigma_n} = \frac{T_n - np}{\sqrt{npq}}$$

  is approximately $N(0, 1)$.

(d) If $T_m \sim Binomial(m,p)$ and $S_n \sim Binomial(n,p)$ are independent, then $T_m + S_n \sim Binomial(m+n,p)$.

(e) Given that $T_n \equiv X_1 + \ldots + X_n = m$, for an integer i with $1 \le i \le n$, $T_i \sim Hypergeometric(i,n,m)$; $(T_i|T_n = m) \sim Hypergeometric(i,n,m)$. See HO #1 for the notation Hypergeometric(R,N,n).

- Negative Binomial facts:

(a) The basic rv Y has mean $\mu_Y = 1/p$ and variance $\sigma_Y^2 = q/p^2$.

(b) The rv W_r has mean $\mu_r = r/p$ and variance $\sigma_r^2 = rq/p^2$.

(c) The rv

$$Z_r \equiv \frac{W_r - \mu_r}{\sigma_r} = \frac{W_r - r/p}{\sqrt{rq/p^2}}$$

is approximately $N(0,1)$.

(d) If $W_r \sim NegativeBinomial(r,p)$ and $W_s \sim NegativeBinomial(s,p)$ are independent, then $W_r + W_s \sim NegativeBinomial(r+s,p)$.

- Geometric facts:

(a) $P(Y > k) = q^k p + q^{k+1}p + q^{k+2}p + \cdots = q^k p/(1-q) = q^k$ for $k = 1,2,\ldots$.

(b) $P(Y > i+k|Y > i) = P(Y > i+k)/P(Y > i) = q^{k+i}/q^i = q^k = P(Y > k)$. This is the *memoryless* property of the Geometric distribution.

- Also

$$P(X_1 = x_1, \ldots, X_n = x_n | X_1 + \cdots + X_n = m) = \begin{cases} 1/\binom{n}{m} & \text{if } x_1 + \cdots x_n = m \\ 0 & \text{if not.} \end{cases}$$

That is, given the number of successes, their location is random and does not depend on p.