

1. Maximum variance Bernoulli and binomial random variables

- (a) The variance of the Bernoulli random variable X is $p(1-p)$ which is zero at each end of the interval of possible values for p , $[0,1]$. The function is positive in between. The derivative of the function is $1-2p$ which is zero precisely when $p = \frac{1}{2}$. So $p = \frac{1}{2}$ maximizes the variance, and the maximum value is $\frac{1}{4}$.
- (b) The variance of the binomial random variable Y is $np(1-p)$, a constant n times larger than the variance in part (a). So by part (a), the variance of Y is also maximized by $p = \frac{1}{2}$, and the maximum variance is $\frac{n}{4}$.

2. Some calculations with the Poisson distribution

- (a) Parameter n is the number of leaves on the tree at the beginning of the one minute observation interval, and p is the probability that any given leaf will fall off during the observation interval.
- (b) The estimator $\hat{\lambda}_{ML}$ is the value of λ that maximizes $P[X = 12] = \frac{e^{-\lambda}\lambda^{12}}{12!}$. Equivalently, it maximizes $e^{-\lambda}\lambda^{12}$, or equivalently it maximizes the natural log of this quantity, $-\lambda + 12\ln(\lambda)$. This expression tends to $-\infty$ as $\lambda \rightarrow 0$ or $\lambda \rightarrow \infty$. The derivative $-1 + \frac{12}{\lambda}$ is zero only when $\lambda = 12$. Thus, $\hat{\lambda}_{ML} = 12$.
- (c)

$$P[X = 2|X \leq 3] = \frac{P[\{X = 2\} \cap \{X \leq 3\}]}{P[X \leq 3]} = \frac{P[\{X = 2\}]}{P[X \leq 3]} = \frac{\lambda^2/2}{1 + \lambda + \frac{\lambda^2}{2} + \frac{\lambda^3}{6}}$$

(The factors $e^{-\lambda}$ cancel out.)

3. On the shape of the Poisson distribution

“**Proposition 6.1'** If X is a Poisson random variable with parameter λ , where $\lambda > 0$, then as k goes from 0 to ∞ , $P\{X = k\}$ first increases monotonically and then decreases monotonically, reaching its largest value when k is the largest integer less than or equal to λ .

Proof: We prove the proposition by considering $P\{X = k\}/P\{X = k-1\}$ and determining for what values of k it is greater than or less than 1. Now,

$$\begin{aligned} \frac{P[\{X = k\}]}{P[\{X = k-1\}]} &= \frac{\frac{e^{-\lambda}\lambda^k}{k!}}{\frac{e^{-\lambda}\lambda^{k-1}}{(k-1)!}} \\ &= \frac{\lambda}{k} \end{aligned}$$

Hence $P\{X = k\} \geq P\{X = k-1\}$ if and only if $k \leq \lambda$, and the proposition is proved.”

(This is essentially Proposition 6.1 and its proof in the limit $n \rightarrow \infty$, $p \rightarrow 0$ with $np = \lambda$ fixed.)

4. Poisson approximation and the birthday paradox

- (a) There are $365 - k$ dates left to choose from, so the probability is $\frac{365-k}{365}$.
- (b) $1 \cdot \frac{364}{365} \cdot \frac{363}{365} \cdots \frac{365-k+1}{365}$

k	$P[\text{all } k \text{ distinct}]$	Poisson approximation
10	0.883052	0.884009
20	0.588562	0.594195
30	0.293684	0.303680

- (c) $P[E_{12}] = \frac{1}{365}$ because given the birthday of the first student, the chances of the second student having the same birthday is $\frac{1}{365}$.

$P[E_{12}E_{13}] = \frac{1}{365^2}$ because given the birthday of the first student, the chances of both the second and third student having the same birthday as the first student is $\frac{1}{365^2}$.

$P[E_{12}E_{13}E_{23}] = \frac{1}{365^3}$ since the event $E_{12}E_{13}E_{23}$ is identical to the event $E_{12}E_{13}$.

Yes, the events E_{ij} are pairwise independent since $P[E_{ij}E_{i'j'}] = \frac{1}{365^2} = P[E_{ij}]P[E_{i'j'}]$ whenever ij and $i'j'$ are distinct pairs with $i < j$ and $i' < j'$.

No, the events are not independent since $P[E_{12}E_{13}E_{23}] = \frac{1}{365^3} \neq P[E_{12}]P[E_{13}]P[E_{23}]$.

- (d) The suggested approximation is $P[X = k] \approx \exp(-\frac{k}{2}/365) = \exp(-\frac{k(k-1)}{2 \times 365})$. Numerical values are indicated in the table.

5. Conditional lifetimes and the memoryless property of the geometric distribution

(a) $P[X > 3] = 1 - p(3) = 0.8$, $P[X > 8|X > 5] = \frac{P[\{X > 8\} \cap \{X > 5\}]}{P[X > 5]} = \frac{P[X > 8]}{P[X > 5]} = \frac{0}{0.60} = 0$.

(So a five year old working battery is not equivalent to a new one!)

(b) $P[Y > 3] = P[\text{miss first three shots}] = (1 - p)^3$. On the other hand,

$$P[Y > 8|Y > 5] = \frac{P[\{Y > 8\} \cap \{Y > 5\}]}{P[Y > 5]} = \frac{P[Y > 8]}{P[Y > 5]} = \frac{(1 - p)^8}{(1 - p)^5} = (1 - p)^3.$$

(A player that has missed five shots is equivalent to a player just starting to take shots.)

(c) Y has a geometric distribution. (Part (b) illustrates the fact that the geometric distribution is the memoryless lifetime distribution on the positive integers.)

6. Packet length choice for noisy channels

(a) $p_{S,packet} = (1 - p_{e,Byte})^{n+h}$, because a packet is successful if and only if all $n + h$ bytes are correctly received.

(b) $E[X] = nP[X = n] = np_{S,packet}$, so $\eta = \frac{n}{n+h}p_{S,packet}$. (The term $\frac{n}{n+h}$ accounts for the loss in efficiency due to the packet headers and the term $p_{S,packet}$ accounts for the fraction of packet transmissions that are successful.)

(c) Very small n is not efficient due to header overhead, reflected in the factor $\frac{n}{n+h}$.

(d) Very large n is not efficient due to very small packet success probability, reflected by the exponent n in the expression for $p_{S,packet}$.

(e) Using a small computer program to compute η for a range of n , we find that $n = 53$ maximizes η , though the value of η is nearly the same for all n near 53. The value $n = 53$ yields $\eta \approx 0.151$.

7. Cumulative distribution functions

(a) $P[X \leq 1] = 0.55$

(b) $P[X > 1] = 1 - P[X \leq 1] = 0.45$

(c) $P[|X| \leq 3] = P[X \leq 3] - P[X \leq -3] = 0.65 - 0.35 = 0.30$.

(d) $P[X^2 \leq 9] = P[|X| \leq 3] = 0.30$. (See previous part.)