Discrete Probability Distributions

Discrete Random Variables: A random variable X is **discrete** if the possible values of X are countable.

If these values are $\{x_1, x_2, \dots, x_k, \dots, x_n\}$ and $p_k = P(X = x_k)$, then $F_X(x) = \sum_{k:x_k \le x} p_k$, $0 \le p_k \le 1$ and $\sum_k p_k = 1$.

Expected Value (or mean): $\boxed{E(X) = \sum_k x_k p_k.}$

Note 1: As the number of trials approach infinity, the expected value approaches the arithmetic mean.

Note 2: Consider a k-dimensional space with a random variable vector $\mathbf{x} = (x_1, \dots, x_n)$ and a probability vector $\mathbf{p} = (p_1, \dots, p_n)$, $\mathrm{E}(X) = \mathbf{x} \cdot \mathbf{p}$.

Theorem: If $Y = g(X)$, $\mathrm{E}(Y) = \mathrm{E}(g(X)) = \sum_k g(x_k) p_k$. (If g is linear, $\mathrm{E}(g(X)) = g(\mathrm{E}(X))$.

Variance: $\boxed{\mathrm{Var}(X) = \mathrm{E}\big((X - \mathrm{E}(X))^2\big) = \sum_k \big(x_k - \mu\big)^2 p_k, \quad \text{where } \mu = \mathrm{E}(X).}$

Note: Variance is "produced" by taking the difference of the random variable and the expected value $\big(X - E(X)\big)$, square it $\big(X - E(X)\big)^2$, then the expected value of this new random variable is the variance.

Standard Deviation : $\boxed{\mathrm{SD}(X) = \sqrt{\mathrm{Var}(X)}.}$

Theorem: $\mathrm{Var}(X) = \mathrm{E}(X^2) - \big(\mathrm{E}(X)\big)^2$.

Proof: Let $\mu = \mathrm{E}(X)$ and $g(X) = (X - \mu)^2$, then $\mathrm{Var}(X) = \sum_k (x_k - \mu)^2 p_k = \sum_k (x_k^2 - 2x_k\mu + \mu^2)\, p_k$

$$= \sum_k x_k^2 p_k - 2\mu \sum_k x_k p_k + \mu^2 \sum_k p_k = \mathrm{E}(X^2) - 2\mu^2 + \mu^2 = \mathrm{E}(X^2) - \big(\mathrm{E}(X)\big)^2.$$

Lemma : $\mathrm{E}\left(\sum_{r=0}^{n} a_r X^r\right) = \sum_{r=0}^{n} a_r \mathrm{E}(X^r)$

Proof: $\mathrm{LHS} = \sum_k \left(\sum_{r=0}^{n} a_r x_k^r\right) p_k = \sum_{r=0}^{n} a_r \sum_k x_k^r p_k = \sum_{r=0}^{n} a_r \mathrm{E}(X^r) = \mathrm{RHS}.$

Theorem: $\mathrm{E}(aX + b) = a\mathrm{E}(X) + b$, $\mathrm{Var}(aX + b) = a^2 \mathrm{Var}(X)$, $\mathrm{SD}(aX + b) = |a| \mathrm{SD}(X)$. ($a$ and b are constants.)

Proof: $\mathrm{E}(aX + b) = a\mathrm{E}(X) + b$ is the result of the previous lemma.

$$\mathrm{Var}(aX + b) = \mathrm{E}\Big(\big((aX + b) - \mathrm{E}(aX + b)\big)^2\Big) = \mathrm{E}\Big(\big(aX + b - a\mathrm{E}(X) - b\big)^2\Big) = \mathrm{E}\big(a^2(X - \mathrm{E}(X))^2\big)$$

$$= a^2 \mathrm{E}\big((X - \mathrm{E}(X))^2\big) = a^2 \mathrm{Var}(X).$$

$$\mathrm{SD}(aX + b) = \sqrt{\mathrm{Var}(aX + b)} = \sqrt{a^2 \mathrm{Var}(X)} = |a| \sqrt{\mathrm{Var}(X)} = |a| \mathrm{SD}(X).$$

The Binomial Distribution: $\boxed{B(n,p,k) = \binom{n}{k} p^k (1-p)^{n-k} \quad \text{for } n \in \mathbb{N} \text{ and } k = 0,1,\ldots,n.}$

This represents the probability of exactly k occurrences, each with probability p, in n trials.

Let $q = 1-p$. $\displaystyle\sum_k B(n,p,k) = \sum_{k=0}^{n} B(n,p,k) = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} = (p+q)^n = 1.$

If X is the random variable that measures the number of outcome of some Bernoulli process with n trials, each with probability p, then X has the binomial distribution $B(n,p)$. We write $X \sim B(n,p)$. ($x_k = k$.)

Note: Not to be confused n trials and $n+1$ possible outcomes $k = 0,1,\ldots,n$, including the case of no outcomes.

$\boxed{\mathrm{E}(X) = np.}$ Proof: $p+q=1$, $\displaystyle\frac{d}{dp}\left(\frac{p}{q}\right) = \frac{1}{q^2}$, $\displaystyle\frac{d}{dp}\left(\frac{p}{q}\right)^k = k\left(\frac{p}{q}\right)^{k-1}\frac{1}{q^2}$, $\displaystyle k\left(\frac{p}{q}\right)^{k-1} = q^2\frac{d}{dp}\left(\frac{p}{q}\right)^k$.

$$\mathrm{E}(X) = \sum_{k=0}^{n} x_k p_k = \sum_{k=0}^{n} k\binom{n}{k} p^k q^{n-k} = pq^{n-1}\sum_{k=0}^{n}\binom{n}{k} k\left(\frac{p}{q}\right)^{k-1} = pq^{n-1}\sum_{k=0}^{n}\binom{n}{k} q^2\frac{d}{dp}\left(\frac{p}{q}\right)^k$$

$$= pq^{n+1}\frac{d}{dp}\left[\sum_{k=0}^{n}\binom{n}{k}\left(\frac{p}{q}\right)^k\right] = pq^{n+1}\frac{d}{dp}\left[1+\frac{p}{q}\right]^n = pq^{n+1}\frac{d}{dp}\left[\frac{1}{q}\right]^n = pq^{n+1}\frac{n}{q^{n+1}} = np.$$

$\boxed{\mathrm{Var}(X) = npq = np(1-p).}$

Proof: $$\frac{d^2}{dp^2}\left(\frac{p}{q}\right)^k = \frac{d}{dp}\left[k\left(\frac{p}{q}\right)^{k-1}\frac{1}{q^2}\right] = \frac{k}{q^2}\frac{d}{dp}\left[\left(\frac{p}{q}\right)^{k-1}\right] + k\left(\frac{p}{q}\right)^{k-1}\frac{d}{dp}\frac{1}{q^2}$$

$$= \frac{k(k-1)}{q^4}\left(\frac{p}{q}\right)^{k-2} + 2k\left(\frac{p}{q}\right)^{k-1}\frac{1}{q^3} = \frac{k^2}{q^4}\left(\frac{p}{q}\right)^{k-2} - \frac{k}{q^4}\left(\frac{p}{q}\right)^{k-2} + 2k\left(\frac{p}{q}\right)^{k-1}\frac{1}{q^3}$$

$$= \frac{k^2}{q^4}\left(\frac{p}{q}\right)^{k-2} - k\left(\frac{p}{q}\right)^{k-1}\left(\frac{1}{q^4}\left(\frac{p}{q}\right)^{-1} - \frac{2}{q^3}\right) = \frac{k^2}{q^4}\left(\frac{p}{q}\right)^{k-2} - q^2\frac{d}{dp}\left(\frac{p}{q}\right)^k\cdot\left(\frac{1}{q^4}\left(\frac{p}{q}\right)^{-1} - \frac{2}{q^3}\right).$$

$$= \frac{k^2}{q^4}\left(\frac{p}{q}\right)^{k-2} - \frac{1}{q^4}\frac{d}{dp}\left(\frac{p}{q}\right)^k\cdot\left(\frac{q^3}{p} - 2q^3\right).$$

$$k^2\left(\frac{p}{q}\right)^{k-2} = q^4\frac{d^2}{dp^2}\left(\frac{p}{q}\right)^k + \frac{d}{dp}\left(\frac{p}{q}\right)^k\cdot\left(\frac{q^3}{p} - 2q^3\right).$$

Let $\mu = \mathrm{E}(X) = np$, $\displaystyle \mathrm{Var}(X) = \mathrm{E}(X^2) - \big(\mathrm{E}(X)\big)^2 = \left[\sum_{k=0}^{n} x_k^2 p_k\right] - (np)^2 = \left[\sum_{k=0}^{n} k^2\binom{n}{k}p^k q^{n-k}\right] - (np)^2$

$$= p^2q^{n-2}\left[\sum_{k=0}^{n}\binom{n}{k}k^2\left(\frac{p}{q}\right)^{k-2}\right] - (np)^2$$

$$= p^2q^{n-2}\left[\sum_{k=0}^{n}\binom{n}{k}\left[q^4\frac{d^2}{dp^2}\left(\frac{p}{q}\right)^k + \frac{d}{dp}\left(\frac{p}{q}\right)^k\cdot\left(\frac{q^3}{p} - 2q^3\right)\right]\right] - (np)^2$$

$$= p^2q^{n-2}\left[q^4\frac{d^2}{dp^2}\sum_{k=0}^{n}\binom{n}{k}\left(\frac{p}{q}\right)^k + \left(\frac{q^3}{p} - 2q^3\right)\cdot\frac{d}{dp}\sum_{k=0}^{n}\binom{n}{k}\left(\frac{p}{q}\right)^k\right] - (np)^2$$

$$= p^2q^{n-2}\left[q^4\frac{d^2}{dp^2}\left(1+\frac{p}{q}\right)^n + \left(\frac{q^3}{p} - 2q^3\right)\cdot\frac{d}{dp}\left(1+\frac{p}{q}\right)^n\right] - (np)^2$$

$$= p^2q^{n-2}\left[q^4\frac{d^2}{dp^2}q^{-n} + \left(\frac{q^3}{p} - 2q^3\right)\cdot\frac{d}{dp}q^{-n}\right] - (np)^2$$

$$= p^2q^{n-2}\left[q^4 n(n+1)q^{-n-2} + \left(\frac{q^3}{p} - 2q^3\right)nq^{-n-1}\right] - (np)^2$$

$$= p^2q^{n-2}q^4n(n+1)q^{-n-2} + p^2q^{n-2}\frac{q^3}{p}nq^{-n-1} - p^2q^{n-2}\cdot 2q^3nq^{-n-1} - (np)^2$$

$$= n(n+1)p^2 + np - 2np^2 - n^2p^2 = n^2p^2 + np^2 + np - 2np^2 - n^2p^2 = np - np^2 = np(1-p).$$

The Geometric Distribution: $\boxed{G(p,k) = (1-p)^{k-1}p \quad \text{for } k = 1, 2, \dots\ .}$

This represents the probability of the first occurrence happening on the kth trial, each trial with probability p.

Let $q = 1 - p$. $\displaystyle\sum_{k=1}^{\infty} G(p,k) = \sum_{k=1}^{\infty} q^{k-1}p = p\sum_{k=0}^{\infty} q^k = p \cdot \frac{1}{1-q} = 1.$

Consider an infinite Bernoulli process of trials each of which has a probability of p. If the random variable X measures the number of trials conducted until the first occurrence, then X has the geometric distribution $G(p)$. We write $X \sim G(p)$. $(x_k = k)$

$$\boxed{\mathrm{E}(X) = \frac{1}{p}.}$$

Proof: $\displaystyle\mathrm{E}(X) = \sum_{k=1}^{\infty} x_k p_k = \sum_{k=1}^{\infty} kq^{k-1}p = \sum_{k=1}^{\infty} \frac{d}{dq}q^k \cdot p = p \cdot \frac{d}{dq}\sum_{k=1}^{\infty} q^k = p \cdot \frac{d}{dq}\left(\frac{q}{1-q}\right) = p \cdot \frac{d}{dq}\left(\frac{q}{p}\right) = p \cdot \frac{p+q}{p^2} = \frac{1}{p}.$

$$\boxed{\mathrm{Var}(X) = \frac{q}{p^2} = \frac{1-p}{p^2}.}$$

Proof: $\displaystyle\frac{d}{dq}q^k = kq^{k-1}, \quad \frac{d^2}{dq^2}q^k = k(k-1)q^{k-2} = (k^2q^{k-1} - kq^{k-1})q^{-1},$

$$k^2q^{k-1} = q\frac{d^2}{dq^2}q^k + kq^{k-1} = q\frac{d^2}{dq^2}q^k + \frac{d}{dq}q^k.$$

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - \big(\mathrm{E}(X)\big)^2 = \left(\sum_{k=1}^{\infty} x_k^2 p_k\right) - \tfrac{1}{p}^2 = \left(\sum_{k=1}^{\infty} k^2q^{k-1}p\right) - p^{-2}$$

$$= p\sum_{k=1}^{\infty}\left(q\frac{d^2}{dq^2}q^k + \frac{d}{dq}q^k\right) - p^{-2} = p\left(q\frac{d^2}{dq^2}\sum_{k=1}^{\infty} q^k + \frac{d}{dq}\sum_{k=1}^{\infty} q^k\right) - p^{-2} = p\left(q\frac{d^2}{dq^2}\left(\frac{q}{1-q}\right) + \frac{d}{dq}\left(\frac{q}{1-q}\right)\right) - p^{-2}$$

$$= p\left(q\frac{d^2}{dq^2}\left(\frac{q}{p}\right) + \frac{d}{dq}\left(\frac{q}{p}\right)\right) - p^{-2} = p\left(q \cdot \frac{2}{p^3} + \frac{1}{p^2}\right) - p^{-2} = \frac{2q}{p^2} + \frac{p}{p^2} - \frac{1}{p^2} = \frac{2q+p-1}{p^2} = \frac{q}{p^2} = \frac{1-p}{p^2}.$$

Theorem: $P(X > n) = (1-p)^n = q^n$, where $X \sim G(p)$, $n = 1, 2, \dots$.

Proof: $\displaystyle P(X > n) = \sum_{k=n+1}^{\infty} q^{k-1}p = pq^n\sum_{k=0}^{\infty} q^k = pq^n \cdot \frac{1}{1-q} = pq^n \cdot \frac{1}{p} = q^n.$