Expected values of sums of RVs

For any RVs $X_1, X_2, \ldots, X_n$ (correlated, independent, whatever...)

$$W = X_1 + X_2 + \cdots + X_n.$$

① What is $E[W]$? $= E[X_1 + X_2 + \cdots + X_n]$

$$= E[X_1] + E[X_2] + \cdots + E[X_n] \quad \text{ALWAYS.}$$

$$= \mu_1 + \mu_2 + \cdots + \mu_n \quad (\text{where } E[X_i] = \mu_i)$$

② What is $Var(W)$?

$$Var(W) = \sum_{i=1}^{n} Var(X_i) + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} Cov(X_i, X_j)$$

Proof: $Var(W) \overset{DEF}{=} E[(W - E[W])^2] = E\left[\left(\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} \mu_i\right)^2\right].$

$$= E\left[\left(\sum_{i=1}^{n} (X_i - \mu_i)\right)\left(\sum_{j=1}^{n} (X_j - \mu_j)\right)\right].$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} Cov(X_i, X_j) \quad (*) \quad \left(E[(X_i - \mu_i)(X_j - \mu_j)] \overset{DEF}{=} Cov(X_i, X_j)\right)$$

Now, recognize
- $Cov(X_i, X_i) = Var(X_i)$
- $Cov(X_i, X_j) = Cov(X_j, X_i)$

} use this to re-write

GENERAL

$$= \sum_{i=1}^{n} Var(X_i) + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} Cov(X_i, X_j)$$

if $X_1, \ldots, X_n$ are ~~independent~~ uncorrelated

$$= \sum_{i=1}^{n} Var(X_i) + 0$$

all terms are 0 when $X_1, \ldots, X_n$ ~~independent~~ uncorrelated

EX: At a party with $n \geq 2$ people, each person throws a hat into a common box. The box is shaken and each person blindly draws a hat at random from the box without replacement. We say we a match occurs if a person draws his/her own hat. Let V_n denote the # of matches. ③

Find $E[V_n]$ and $var[V_n]$:

$V_n \in [0, 1, \ldots, n]$.

Let X_i denote an indicator random variable (Bernouilli) such that

$$X_i = \begin{cases} 1 & \text{if person } i \text{ picks own hat} \\ 0 & \text{else} \end{cases}$$

USEFUL TRICK!

Then $V_n = X_1 + X_2 + \cdots + X_n$.

Are X_i independent? If $n=2$ clearly if we know X_1 then we know X_2.

So NO, not independent.

Note that $P(X_i = 1) = 1/n$. (chance of drawing own hat is $1/n$).

hence $E[X_i] = 1 \cdot P(X_i = 1) + 0 \cdot P(X_i = 0) = \frac{1}{n}$

Then $E[V_n] \overset{\text{ALWAYS.}}{=} E[X_1] + \cdots + E[X_n] = \frac{1}{n} + \frac{1}{n} + \cdots + \frac{1}{n} = 1$

To get $var(V_n)$ we need $Var(X_i)$ and $cov(X_i, X_j)$.

$Var(X_i) = E[X_i^2] - (E[X_i])^2 = (1)^2 P(X_i = 1) + (0)^2 P(X_i = 0) = \frac{1}{n}$ $-\left(\frac{1}{n}\right)^2$

$= \frac{1}{n} - \frac{1}{n^2}$

$$\mathrm{Cov}(X_i, X_j) \overset{\text{DEF}}{=} E[X_i X_j] - \underbrace{E[X_i]}_{\frac{1}{n}}\,\underbrace{E[X_j]}_{\frac{1}{n}}.$$

$$= P(X_i=0, X_j=0)\,0\cdot 0 + P(X_i=0, X_j=1)\,0\cdot 1 + P(X_i=1, X_j=0)\,1\cdot 0$$

$$+ P(X_i=1, X_j=1)\,1\cdot 1 - \frac{1}{n^2}$$

<u>Only need $P(X_i=1, X_j=1)$</u> $= P(X_i=1 \mid X_j=1)\,P(X_j=1).$

$$= \frac{1}{n-1}\cdot\frac{1}{n}.$$

So, $\mathrm{Cov}(X_i, X_j) = \frac{1}{n-1}\cdot\frac{1}{n} - \frac{1}{n}\cdot\frac{1}{n}$

$(V_n = X_1 + X_2 + \cdots + X_n)$

$$\mathrm{Var}(V_n) = \mathrm{Var}(X_1) + \cdots + \mathrm{Var}(X_n) + 2\,\mathrm{Cov}(X_1, X_2) + 2\,\mathrm{Cov}(X_1, X_3) + \cdots + 2\,\mathrm{Cov}(X_1, X_n)$$

$$+ \cdots + 2\,\mathrm{Cov}(X_{n-1}, X_n).$$

$$= n\,\mathrm{Var}[X_i] + n(n-1)\,\mathrm{Cov}(X_i, X_j) \quad (\text{for } i \neq j)$$

$$= n\left[\frac{1}{n} - \frac{1}{n^2}\right] + n(n-1)\left[\frac{1}{n-1}\cdot\frac{1}{n} - \frac{1}{n}\cdot\frac{1}{n}\right]$$

$$= 1$$