Lecture 11

0.1 Variance of Common Distributions

0.1.1 Geometric Distribution

Recall the definition of the Geometric distribution: $X \sim \text{Geom}(p)$ if

$$\mathbb{P}(X = k) = (1-p)^{k-1} \cdot p,$$

for every positive integer k.

Claim 0.1. $\text{Var}(X) = \frac{1-p}{p^2}$.

Proof. We start with calculating $\mathbb{E}(X^2)$.

$$\begin{aligned}
\mathbb{E}\left(X^2\right) &= \sum_{k=1}^{\infty} k^2 \cdot \mathbb{P}(X = k) \\
&= \sum_{k=1}^{\infty}(k^2 - 2k + 1) \cdot (1-p)^{k-1} \cdot p + \sum_{k=1}^{\infty} 2k \cdot (1-p)^{k-1} \cdot p - \sum_{k=1}^{\infty}(1-p)^{k-1} \cdot p \\
&= \sum_{k=1}^{\infty}(k-1)^2 \cdot (1-p)^{k-1} \cdot p + 2\mathbb{E}(X) - 1 \\
&= \sum_{m=0}^{\infty} m^2 \cdot (1-p)^m \cdot p + 2\mathbb{E}(X) - 1 \\
&= (1-p) \cdot \sum_{m=1}^{\infty} m^2 \cdot (1-p)^{m-1} \cdot p + \frac{2}{p} - 1 \\
&= (1-p) \cdot \mathbb{E}\left(X^2\right) + \frac{2-p}{p},
\end{aligned}$$

where the fourth equality holds by the substitution $m = k - 1$ and the fifth equality holds since, as was shown in Lecture 9, $\mathbb{E}(X) = 1/p$. A straightforward calculation then yields $\mathbb{E}(X^2) = \frac{2-p}{p^2}$. This in turn implies that

$$\text{Var}(X) = \mathbb{E}\left(X^2\right) - (\mathbb{E}(X))^2 = \frac{2-p}{p^2} - \left(\frac{1}{p}\right)^2 = \frac{1-p}{p^2}.$$

$\square$

0.1.2 Negative-Binomial Distribution

Recall the definition of the Negative-Binomial distribution: $X \sim \mathrm{NB}\left(r, p\right)$ if

$$\mathbb{P}\left(X = n\right) = \binom{n-1}{r-1} \cdot p^r \cdot (1-p)^{n-r},$$

for every integer $n \geq r$.

Claim 0.2. $\mathrm{Var}\left(X\right) = \frac{r(1-p)}{p^2}$.

Proof. We first calculate $\mathbb{E}\left(X^2\right)$. Observe that

$$\begin{aligned}
\mathbb{E}\left(X(X+1)\right) &= \sum_{n=r}^{\infty} n(n+1) \cdot \binom{n-1}{r-1} \cdot p^r \cdot (1-p)^{n-r} \\
&= r(r+1) \cdot \sum_{n=r}^{\infty} \binom{n+1}{r+1} \cdot p^r \cdot (1-p)^{n-r} \\
&= \frac{r(r+1)}{p^2} \cdot \sum_{n=r}^{\infty} \binom{n+1}{r+1} \cdot p^{r+2} \cdot (1-p)^{n-r} \\
&= \frac{r(r+1)}{p^2} \cdot \sum_{m=r+2}^{\infty} \binom{m-1}{(r+2)-1} \cdot p^{r+2} \cdot (1-p)^{m-(r+2)} \\
&= \frac{r^2+r}{p^2},
\end{aligned}$$

where the second equality follows from the identity $\binom{a}{b} = \frac{a}{b} \cdot \binom{a-1}{b-1} = \frac{a(a-1)}{b(b-1)} \cdot \binom{a-2}{b-2}$, the fourth equality holds by the substitution $m = n+2$, and the last equality holds since its left hand side is the sum of probabilities of a random variable $Y \sim \mathrm{NB}\left(r+2, p\right)$ over the support of its distribution. Therefore, it follows by the linearity of expectation that

$$\mathbb{E}\left(X^2\right) = \mathbb{E}\left(X(X+1)\right) - \mathbb{E}\left(X\right) = \frac{r^2+r}{p^2} - \frac{r}{p}.$$

We conclude that

$$\mathrm{Var}\left(X\right) = \mathbb{E}\left(X^2\right) - \left(\mathbb{E}\left(X\right)\right)^2 = \frac{r^2+r}{p^2} - \frac{r}{p} - \frac{r^2}{p^2} = \frac{r(1-p)}{p^2}.$$

$\square$

0.1.3 Hypergeometric Distribution

Recall the definition of the Hypergeometric distribution: $X \sim \mathrm{Hyp}\left(N, D, n\right)$ if

$$\mathbb{P}\left(X = k\right) = \frac{\binom{D}{k} \cdot \binom{N-D}{n-k}}{\binom{N}{n}},$$

for every integer $0 \leq k \leq n$.

Claim 0.3. $\mathrm{Var}\left(X\right) = \frac{D \cdot n \cdot (N-D) \cdot (N-n)}{N^2 \cdot (N-1)}$.

The proof will be shown in the practical session.

0.1.4 Poisson Distribution

Recall the definition of the Poisson distribution: $X \sim \text{Poi}(\lambda)$ if

$$\mathbb{P}(X = k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!},$$

for every non-negative integer k.

Claim 0.4. $\text{Var}(X) = \lambda$.

The proof will be shown in the practical session.

0.2 Covariance

Throughout this subsection, let X and Y be two random variables with finite expectation.

Definition 0.5 (Covariance)**.** *The* Covariance *of X and Y is defined to be*

$$\text{Cov}(X, Y) = \mathbb{E}((X - \mathbb{E}(X)) \cdot (Y - \mathbb{E}(Y))).$$

In particular,

$$\text{Cov}(X, X) = \text{Var}(X).$$

Claim 0.6. *If X and Y have finite variance, then $\mathbb{E}(|X \cdot Y|)$, $\mathbb{E}((X-Y)^2)$, and $\mathbb{E}((X+Y)^2)$ are finite as well.*

Proof. Observe that $(X+Y)^2 \leq 2(X^2+Y^2)$. Therefore,

$$\mathbb{E}\left((X+Y)^2\right) \leq \mathbb{E}\left(2(X^2+Y^2)\right) = 2\mathbb{E}\left(X^2\right) + 2\mathbb{E}\left(Y^2\right),$$

where the inequality above holds by the monotonicity of expectation, and the equality holds by the linearity of expectation. Since $\text{Var}(X)$ and $\text{Var}(Y)$ are finite by assumption, $\mathbb{E}(X^2)$ and $\mathbb{E}(Y^2)$ are finite as well, implying that $\mathbb{E}((X+Y)^2)$ is finite. An analogous argument shows that $\mathbb{E}((X-Y)^2)$ is finite as well. Next, observe that

$$\begin{aligned} |X \cdot Y| &= \frac{1}{4} \cdot \left|(X+Y)^2 - (X-Y)^2\right| \leq \frac{1}{4} \cdot \left(\left|(X+Y)^2\right| + \left|(X-Y)^2\right|\right) \\ &= \frac{1}{4} \cdot \left((X+Y)^2 + (X-Y)^2\right). \end{aligned}$$

Hence $\mathbb{E}(|X \cdot Y|)$ is finite. □

Corollary 0.7. *If X and Y have finite variance, then $\text{Cov}(X, Y)$ is finite. Moreover it holds that*

$$\text{Cov}(X, Y) = \mathbb{E}(X \cdot Y) - \mathbb{E}(X) \cdot \mathbb{E}(Y).$$

Proof. By the previous claim it holds that $\mathbb{E}(X \cdot Y)$ is finite (e.g., since $-|X \cdot Y| \leq X \cdot Y \leq |X \cdot Y|$). Therefore

$$\begin{aligned}
\operatorname{Cov}(X,Y) &= \mathbb{E}\left((X-\mathbb{E}(X)) \cdot (Y-\mathbb{E}(Y))\right) \\
&= \mathbb{E}\left(X \cdot Y - \mathbb{E}(X) \cdot Y - X \cdot \mathbb{E}(Y) + \mathbb{E}(X) \cdot \mathbb{E}(Y)\right) \\
&= \mathbb{E}(X \cdot Y) - \mathbb{E}(X) \cdot \mathbb{E}(Y) - \mathbb{E}(X) \cdot \mathbb{E}(Y) + \mathbb{E}(X) \cdot \mathbb{E}(Y) \\
&= \mathbb{E}(X \cdot Y) - \mathbb{E}(X) \cdot \mathbb{E}(Y).
\end{aligned}$$

□

Definition 0.8 (Correlation)**.** *We say that* X *and* Y *are* positively correlated *if* $\operatorname{Cov}(X,Y) > 0$*, are* negatively correlated *if* $\operatorname{Cov}(X,Y) < 0$*, and are* uncorrelated *if* $\operatorname{Cov}(X,Y) = 0$*.*

Example 1: Let A and B be events in a probability space $(\Omega, \mathbb{P})$ such that $\mathbb{P}(B) > 0$. Let 1_A and 1_B be the indicators of A and B, respectively. Then

$$\begin{aligned}
Cov(1_A, 1_B) &= \mathbb{E}(1_A \cdot 1_B) - \mathbb{E}(1_A) \cdot \mathbb{E}(1_B) = \mathbb{P}(1_A = 1, 1_B = 1) - \mathbb{P}(1_A = 1) \cdot \mathbb{P}(1_B = 1) \\
&= \mathbb{P}(A \cap B) - \mathbb{P}(A) \cdot \mathbb{P}(B) = \mathbb{P}(A|B) \cdot \mathbb{P}(B) - \mathbb{P}(A) \cdot \mathbb{P}(B) \\
&= \mathbb{P}(B)\left(\mathbb{P}(A|B) - \mathbb{P}(A)\right).
\end{aligned}$$

We conclude that 1_A and 1_B are positively correlated if and only if $\mathbb{P}(A|B) > \mathbb{P}(A)$ (that is, conditioning on B increases the probability of A), are negatively correlated if and only if $\mathbb{P}(A|B) < \mathbb{P}(A)$ (that is, conditioning on B decreases the probability of A), and are uncorrelated if and only if $\mathbb{P}(A|B) = \mathbb{P}(A)$ (that is, A and B are independent).

Claim 0.9. *If* X *and* Y *are independent, then* $\operatorname{Cov}(X,Y) = 0$*, i.e.,* X *and* Y *are* uncorrelated*. The converse is not necessarily true.*

Proof. Assume first that X and Y are independent. Then

$$\begin{aligned}
\mathbb{E}(X \cdot Y) &= \sum_z z \cdot \mathbb{P}(XY = z) \\
&= \sum_z z \cdot \sum_{\substack{x,y \\ xy=z}} \mathbb{P}(X = x, Y = y) \\
&= \sum_{x,y} xy \cdot \mathbb{P}(X = x, Y = y) \\
&= \sum_{x,y} x \cdot \mathbb{P}(X = x) \cdot y \cdot \mathbb{P}(Y = y) \\
&= \left(\sum_x x \cdot \mathbb{P}(X = x)\right) \cdot \left(\sum_y y \cdot \mathbb{P}(Y = y)\right) \\
&= \mathbb{E}(X) \cdot \mathbb{E}(Y),
\end{aligned}$$

where the fourth equality holds since X and Y are independent. Hence $Cov(X,Y) = \mathbb{E}(X \cdot Y) - \mathbb{E}(X) \cdot \mathbb{E}(Y) = 0$, i.e., X and Y are uncorrelated.

Next, we give an example of random variables X and Y that are uncorrelated but not independent. Let

$$X \sim \begin{cases} 1 & \frac{1}{3} \\ 0 & \frac{1}{3} \\ -1 & \frac{1}{3} \end{cases}$$

and let $Y = X^2$. Observe that

$$Y \sim \begin{cases} 1 & \frac{2}{3} \\ 0 & \frac{1}{3} \end{cases}$$

It is evident that X and Y are not independent, for example

$$\mathbb{P}(X = 0, Y = 0) = \mathbb{P}(X = 0) = 1/3 \neq 1/9 = 1/3 \cdot 1/3 = \mathbb{P}(X = 0) \cdot \mathbb{P}(Y = 0).$$

Now, observe that

$$\mathbb{E}(X) = 1 \cdot 1/3 + 0 \cdot 1/3 + (-1) \cdot 1/3 = 0$$

and, similarly,

$$\mathbb{E}(XY) = \mathbb{E}\left(X^3\right) = 1^3 \cdot 1/3 + 0^3 \cdot 1/3 + (-1)^3 \cdot 1/3 = 0.$$

Therefore

$$Cov(X,Y) = \mathbb{E}(X \cdot Y) - \mathbb{E}(X) \cdot \mathbb{E}(Y) = \mathbb{E}\left(X^3\right) - \mathbb{E}(X) \cdot \mathbb{E}\left(X^2\right) = 0 - 0 = 0,$$

that is, X and Y are uncorrelated. □

0.3 Properties of covariance

In the following claims, let X, Y and Z be random variables and let a and b be real numbers.

Claim 0.10. $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$.

Proof. $\mathrm{Cov}(X,Y) = \mathbb{E}(X \cdot Y) - \mathbb{E}(X) \cdot \mathbb{E}(Y) = \mathbb{E}(Y \cdot X) - \mathbb{E}(Y) \cdot \mathbb{E}(X) = \mathrm{Cov}(Y,X)$. □

Claim 0.11. $\mathrm{Cov}(aX, bY) = ab \cdot \mathrm{Cov}(X,Y)$.

Proof. We have

$$\begin{aligned} \mathrm{Cov}(aX, bY) &= \mathbb{E}(aX \cdot bY) - \mathbb{E}(aX) \cdot \mathbb{E}(bY) = ab \cdot \mathbb{E}(X \cdot Y) - a\mathbb{E}(X) \cdot b\mathbb{E}(Y) \\ &= ab \cdot [\mathbb{E}(X \cdot Y) - \mathbb{E}(X) \cdot \mathbb{E}(Y)] = ab \cdot \mathrm{Cov}(X,Y). \end{aligned}$$

□

Claim 0.12. $\mathrm{Cov}(X + Z, Y) = \mathrm{Cov}(X,Y) + \mathrm{Cov}(Z,Y)$.

Proof. We have

$$\begin{aligned}\operatorname{Cov}(X+Z,Y) &= \mathbb{E}((X+Z)\cdot Y) - \mathbb{E}(X+Z)\cdot\mathbb{E}(Y)\\ &= \mathbb{E}(X\cdot Y + Z\cdot Y) - (\mathbb{E}(X)+\mathbb{E}(Z))\cdot\mathbb{E}(Y)\\ &= \mathbb{E}(X\cdot Y) - \mathbb{E}(X)\cdot\mathbb{E}(Y) + \mathbb{E}(Z\cdot Y) - \mathbb{E}(Z)\cdot\mathbb{E}(Y)\\ &= \operatorname{Cov}(X,Y) + \operatorname{Cov}(Z,Y).\end{aligned}$$

□

0.4 Variance of Sum

Theorem 0.13. *If X and Y have finite variance, then*

$$\operatorname{Var}(X+Y) = \operatorname{Var}(X) + \operatorname{Var}(Y) + 2\operatorname{Cov}(X,Y).$$

More generally, if $X_1,\ldots,X_n$ have finite variance, then

$$\operatorname{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n}\operatorname{Var}(X_i) + 2\sum_{1\le i<j\le n}\operatorname{Cov}(X_i,X_j).$$

Proof. Let $X = \sum_{i=1}^{n} X_i$. Then

$$\begin{aligned}\operatorname{Var}(X) &= \mathbb{E}\left(X^2\right) - (\mathbb{E}(X))^2\\ &= \mathbb{E}\left(\left(\sum_{i=1}^{n} X_i\right)^2\right) - \left(\mathbb{E}\left(\sum_{i=1}^{n} X_i\right)\right)^2\\ &= \mathbb{E}\left(\sum_{i=1}^{n}\sum_{j=1}^{n} X_iX_j\right) - \left(\sum_{i=1}^{n}\mathbb{E}(X_i)\right)^2\\ &= \sum_{i=1}^{n}\sum_{j=1}^{n}\mathbb{E}(X_iX_j) - \sum_{i=1}^{n}\sum_{j=1}^{n}\mathbb{E}(X_i)\cdot\mathbb{E}(X_j)\\ &= \sum_{i=1}^{n}\left(\mathbb{E}\left(X_i^2\right) - [\mathbb{E}(X_i)]^2\right) + \sum_{1\le i\ne j\le n}(\mathbb{E}(X_iX_j) - \mathbb{E}(X_i)\cdot\mathbb{E}(X_j))\\ &= \sum_{i=1}^{n}\operatorname{Var}(X_i) + \sum_{1\le i\ne j\le n}\operatorname{Cov}(X_i,X_j)\\ &= \sum_{i=1}^{n}\operatorname{Var}(X_i) + 2\sum_{1\le i<j\le n}\operatorname{Cov}(X_i,X_j),\end{aligned}$$

where the last equality holds by Claim 0.10. □

Since, by Claim 0.9, $Cov(X,Y) = 0$ whenever X and Y are independent, the following is an immediate consequence of Theorem 0.13.

Corollary 0.14. *If $X_1,\ldots,X_n$ have finite variance, and they are pairwise independent, then*

$$\operatorname{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n}\operatorname{Var}(X_i).$$