- Outline

Discrete Distributions

X is a *discrete random variable* if it has a *probability mass function* p_k,

$$p_k = \mathrm{P}[X = k]$$
$$0 \le p_k \le 1$$
$$\sum_k p_k = 1$$

Example: If X is a Poisson random variable with mean λ then

$$p_k = e^{-\lambda} \frac{\lambda^k}{k!} \quad k = 0, 1, 2, \ldots$$

For us, discrete random variables will usually be counting variables, meaning the possible values are $\{0, 1, 2, \ldots\}$

Empirical Distribution

Suppose we observe 5 losses with amounts of:

$$500 \quad 600 \quad 650 \quad 650 \quad 750$$

If X is distributed according to the *empirical distribution* then

$$\mathrm{P}[X = 500] = \frac{1}{5} \quad \mathrm{P}[X = 650] = \frac{2}{5}$$

More generally, if we have n data points,

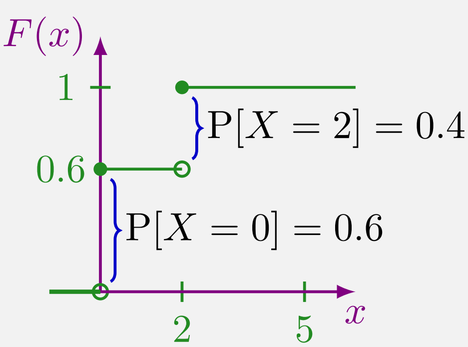
$$\mathrm{P}[X = x] = \frac{\#\text{ of data points} = x}{\text{total }\#\text{ of data points}} = \frac{\#\text{ of data points} = x}{n}$$

Cumulative Distribution Function (CDF)

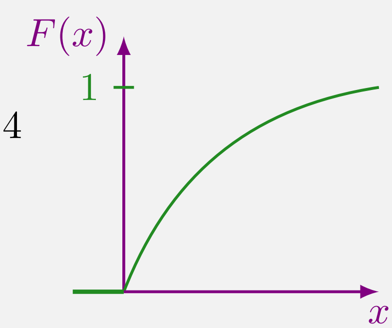
The cumulative distribution function (CDF) of X is

$$F(x) = \mathrm{P}[X \le x]$$
$$F(\infty) = \lim_{x\to\infty} \mathrm{P}[X \le x] = 1$$
$$F(-\infty) = 0$$

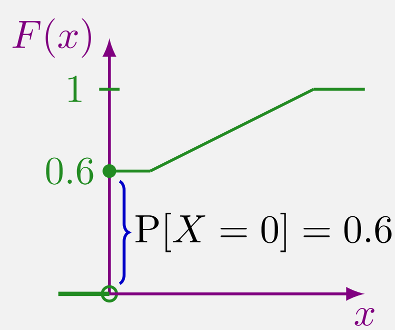
X is a *loss variable* if $X \ge 0$, in which case $F(x) = 0$ for all $x < 0$.



Discrete



Continuous



Mixed

Continuous Distributions

X is a continuous variable if $F(x)$ is continuous. For us, this will also imply that $F(x)$ is differentiable. The density $f(x)$ is given by

$$\begin{aligned} f(x) &= F'(x) \\ \mathrm{P}[a < X \le b] &= \mathrm{P}[X \le b] - \mathrm{P}[X \le a] = F(b) - F(a) \\ &= \int_a^b f(x)dx \\ 1 &= \int_{-\infty}^{\infty} f(x)dx \\ f(x)\,dx &\text{``="}\ \mathrm{P}[x < X \le x + dx] \\ 0 &\le f(x) \end{aligned}$$

There is no upper limit on $f(x)$. In particular, $f(x)$ is not a probability, and can be greater than 1. For example, if X is uniform on $(0, 0.1)$ then $f(x) = 1/(0.1 - 0) = 10$

Survival Function

$$\begin{aligned} S(x) &= \mathrm{P}[X > x] = 1 - F(x) = \text{survival function} \\ S(x) &= e^{-H(x)} \\ H(x) &= -\ln[S(x)] = \text{cumulative hazard function} \\ h(x) &= H'(x) = \text{hazard rate} \\ &= \frac{-S'(x)}{S(x)} = \frac{f(x)}{S(x)} \\ h(x)\,dx &= \frac{f(x)\,dx}{S(x)} \\ &\text{``="}\ \mathrm{P}[x < X \le x + dx \mid X > x] \end{aligned}$$

Percentiles

If X is continuous, then $100p^{th}$ percentile $= \pi_p(X)$
i.e., $F(\pi_p) = \mathrm{P}[X \le \pi_p] = p = p \times 100\%$
$\pi_p(X)$ is denoted $\mathrm{VaR}_p(X)$ on tables.

If X is discrete or mixed, the CDF may jump over p or $F(x) = p$ may not have a unique solution.

In words, if $F(x)$ jumps over p, then π_p is where the jump occurs.
If $F(x) = p$ doesn't have a unique solution, then all x-values with $F(x) = p$ are $100p^{th}$ percentiles, as is the right hand end point of that set.
In equations, π_p is a $100p^{th}$ percentile if

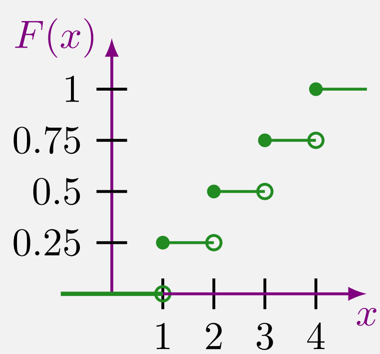
$$\mathrm{P}[X < \pi_p] \le p \le \mathrm{P}[X \le \pi_p]$$
$$\lim_{x \uparrow \pi_p} F(x) = F(\pi_p^-) \le p \le F(\pi_p)$$

This definition has never been tested.

Example

Let X be uniform on $\{1, 2, 3, 4\}$. Find the 10th, 30th, 40th, and 50th percentiles.

$$F(x) = \begin{cases} 0 & x < 1 \\ 0.25 & 1 \le x < 2 \\ 0.5 & 2 \le x < 3 \\ 0.75 & 3 \le x < 4 \\ 1 & 4 \le x \end{cases}$$



$F(x)$ jumps from 0 to 0.25 at 1, so $\pi_{0.1}(X) = 1$.
$F(x)$ jumps from 0.25 to 0.5 at 2, so $\pi_{0.3}(X) = \pi_{0.4}(X) = 2$.
$F(x) = 0.5$ for $2 \le x < 3$.
$F(x^-) \le 0.5$ for $x \le 3$, and $F(x) \ge 0.5$ for $x \ge 2$, so every point with $2 \le x \le 3$ is a 50th percentile.

Example

Suppose that $f(x) = cx^2$ for $0 < x < 5$ and is 0 otherwise. Find the 90th percentile of X.

$$\begin{aligned}
1 &= \int_0^5 cx^2 dx \\
&= c\frac{x^3}{3}\bigg|_0^5 = c \cdot \frac{125}{3} \\
c &= \frac{3}{125} = \frac{1}{\int\limits_0^5 x^2 dx} \\
0.90 &= \int_0^{\pi_{0.9}} \frac{3}{125} x^2 dx \\
&= \frac{1}{125}\pi_{0.9}^3 \\
\pi_{0.9} &= \boxed{4.83}
\end{aligned}$$

Exercise 1

Find the density of X if the hazard rate is

$$h(x) = \begin{cases} 0 & x < 1 \\ 3 & 1 < x < 5 \\ 2x & 5 < x \end{cases}$$

Exercise 1

Find the density of X if the hazard rate is

$$h(x) = \begin{cases} 0 & x < 1 \\ 3 & 1 < x < 5 \\ 2x & 5 < x \end{cases}$$

$$\begin{aligned} H(x) &= 0 \quad x < 1 \\ H(x) &= \int_1^x h(t)\,dt = \int_1^x 3\,dt = 3(x-1) \quad 1 < x < 5 \\ H(x) &= \int_{-\infty}^x h(t)\,dt \quad 5 < x \\ &= H(5) + \int_5^x 2t\,dt \quad 5 < x \\ &= 3 \cdot (5-1) + (x^2 - 25) \quad 5 < x \\ &= x^2 - 13 \quad 5 < x \end{aligned}$$

Exercise 1 (Continued)

$$\begin{aligned} H(x) &= \begin{cases} 0 & x < 1 \\ 3x - 3 & 1 < x < 5 \\ x^2 - 13 & 5 < x \end{cases} \\ S(x) &= e^{-H(x)} \\ &= \begin{cases} 1 & x < 1 \\ e^{-3x+3} & 1 < x < 5 \\ e^{-x^2+13} & 5 < x \end{cases} \\ f(x) &= -S'(x) \\ &= \begin{cases} 0 & x < 1 \\ 3e^{-3x+3} & 1 < x < 5 \\ 2xe^{-x^2+13} & 5 < x \end{cases} \end{aligned}$$

Exercise 2

The cdf of X is $F(x) = 1 - \left(\frac{5}{x+5}\right)^3$ for $x > 0$. Find the hazard rate of x for $x > 0$.

Exercise 2

The cdf of X is $F(x) = 1 - \left(\frac{5}{x+5}\right)^3$ for $x > 0$. Find the hazard rate of x for $x > 0$.

$$
\begin{aligned}
S(x) &= \left(\frac{5}{x+5}\right)^3 \\
H(x) &= -\ln[S(x)] \\
&= -\ln\left[\left(\frac{5}{x+5}\right)^3\right] \\
&= -3\ln(5) + 3\ln(x+5) \\
h(x) &= 0 + \frac{3}{x+5}
\end{aligned}
$$

- Outline

Moments

If X is discrete:

$$\mathrm{E}[X] = \sum_x x \cdot \mathrm{P}[X = x]$$

$$\mathrm{E}\left[X^k\right] = \sum_x x^k \cdot \mathrm{P}[X = x]$$

$$\mathrm{E}[g(X)] = \sum_x g(x) \cdot \mathrm{P}[X = x]$$

If X is continuous

$$\mathrm{E}[X] = \int x \cdot f(x)dx$$

$$\mathrm{E}\left[X^k\right] = \int x^k \cdot f(x)dx$$

$$\mathrm{E}[g(X)] = \int g(x) \cdot f(x)dx$$

For a mixed distribution:

- Use discrete formula over discrete values
- Use continuous formula over continuous values
- Sum the two pieces

Survival Function Approach

Recall that $S(x) = \mathrm{P}[X > x]$.
For any loss variable X, integration by parts gives

$$\begin{aligned} \mathrm{E}[X] &= \int_0^\infty x \cdot f(x)\,dx \\ &= x \cdot [-S(x)]\Big|_0^\infty + \int_0^\infty S(x)dx \\ &= 0 + \int_0^\infty S(x)dx \end{aligned}$$

Also true, but less useful: If $g(0) = 0$,

$$\mathrm{E}[g(X)] = \int_0^\infty g'(x) \cdot S(x)dx$$

Central Moments

$$\begin{aligned} \mathrm{E}\left[X^k\right] &= \mu'_k = k\text{-th } \textit{raw} \text{ moment.} \\ \mu'_1 &= \mu = \mathrm{E}[X] \\ \mathrm{E}\left[(X-\mu)^k\right] &= \mu_k = k\text{-th } \textit{central} \text{ moment} \\ \mu_2 &= \sigma^2 = \mathrm{Var}[X] \\ \mathrm{Var}[X] &= \mathrm{E}\left[(X-\mu)^2\right] = \mathrm{E}\left[X^2\right] - (\mathrm{E}[X])^2 \end{aligned}$$

Bernoulli shortcut: If $\mathrm{P}[X = a] + \mathrm{P}[X = b] = 1$ then

$$\mathrm{Var}[X] = (b-a)^2 \cdot \mathrm{P}[X = a] \cdot \mathrm{P}[X = b]$$

$$\begin{aligned} \text{If } Y = aX \text{ then } \mathrm{E}[Y^k] &= \mathrm{E}[(aX)^k] = a^k \mathrm{E}[X^k] \\ \mu'_k(Y) &= a^k\, \mu'_k(X) \\ \mu_k(Y) &= a^k\, \mu_k(X) \end{aligned}$$

Converting Raw to Central Moments

$$
\begin{aligned}
\mathrm{E}\left[(X-\mu)^2\right] &= \mathrm{Var}[X] = \sigma^2 \\
&= \mathrm{E}\left[X^2 - 2\mu X + \mu^2\right] = \mathrm{E}\left[X^2\right] - 2\mu\mathrm{E}[X] + \mu^2 \\
&= \mathrm{E}\left[X^2\right] - 2\mu \cdot \mu + \mu^2 = \mathrm{E}\left[X^2\right] - \mu^2 \\
\mathrm{E}\left[(X-\mu)^3\right] &= \mathrm{E}\left[X^3 - 3\mu X^2 + 3\mu^2 X - \mu^3\right] \\
&= \mathrm{E}\left[X^3\right] - 3\mu\mathrm{E}\left[X^2\right] + 3\mu^2 \cdot \mathrm{E}[X] - \mu^3 \\
&= \mathrm{E}\left[X^3\right] - 3\mu\mathrm{E}\left[X^2\right] + 3\mu^2 \cdot \mu - \mu^3 \\
&= \mathrm{E}\left[X^3\right] - 3\mu\mathrm{E}\left[X^2\right] + 2\mu^3
\end{aligned}
$$

You can do a similar thing with any central moment. This is rarely tested – remember the process, don't memorize the results!

Coefficient of Variation

$\mu = \mathrm{E}[X] \quad \sigma^2 = \mathrm{Var}[X] \quad \sigma = \mathrm{SD}[X]$

Definition (Coefficient of Variation)

$\mathrm{CV}[X] = \dfrac{\sigma}{\mu}$

Note that for $c > 0$,

$$
\begin{aligned}
\mathrm{CV}[X] &= \frac{\sigma}{\mu} = \frac{\mathrm{SD}[X]}{\mathrm{E}[X]} \\
\mathrm{CV}[cX] &= \frac{\mathrm{SD}[cX]}{\mathrm{E}[cX]} = \frac{c\mathrm{SD}[X]}{c\mathrm{E}[X]} \\
&= \mathrm{CV}[X]
\end{aligned}
$$

Skewness and Kurtosis

Definition (Skewness)

$$\text{Skewness}(X) = \frac{\mathrm{E}\left[(X-\mu)^3\right]}{\sigma^3}$$

Definition (Kurtosis)

$$\text{Kurtosis}(X) = \frac{\mathrm{E}\left[(X-\mu)^4\right]}{\sigma^4}$$

The powers in numerator and denominator match, so for $c > 0$, Skewness(cX) = Skewness(X) and Kurtosis(cX)=Kurtosis(X).

X is symmetric $\Rightarrow$ odd central moments are 0.
Skewness(X)=0 for any symmetric (and thus non-skewed) distribution.

Exercise 1

$\mathrm{P}[X = 100] = 0.3$ and $\mathrm{P}[X = 300] = 0.7$. Find the skewness of X.

Exercise 1

P$[X = 100] = 0.3$ and P$[X = 300] = 0.7$. Find the skewness of X.

$$\begin{aligned}
\mathrm{E}[X] &= 0.3 \cdot 100 + 0.7 \cdot 300 = 240 \\
\mathrm{Var}[X] &= 0.3 \cdot (100 - 240)^2 + 0.7 \cdot (300 - 240)^2 \\
&= (300 - 100)^2 \cdot 0.3 \cdot 0.7 = 8{,}400 \\
\mathrm{E}[(X - \mu)^3] &= 0.3 \cdot (100 - 240)^3 + 0.7 \cdot (300 - 240)^3 \\
&= -672{,}000 \\
\mathrm{Skew}[X] &= \frac{\mathrm{E}[(X - \mu)^3]}{\sigma^3} \\
&= \frac{-672{,}000}{8{,}400^{3/2}} \\
&= \boxed{-0.873}
\end{aligned}$$

Exercise 2

X has hazard rate $h(x) = 3/x$ for $x > 4$ and 0 for $x \leq 4$. Find E$[X]$.

Exercise 2

X has hazard rate $h(x) = 3/x$ for $x > 4$ and 0 for $x \leq 4$. Find $\mathrm{E}[X]$.

$$\begin{aligned} H(x) &= 0 \quad x \leq 4 \\ &= \int_{-\infty}^{x} h(t)\,dt = H(4) + \int_{4}^{x} h(t)\,dt \quad x > 4 \\ &= 0 + 3\ln(t)\Big|_4^x = \ln[(x/4)^3] \quad x > 4 \end{aligned}$$

$$S(x) = \begin{cases} 1 & x \leq 4 \\ \dfrac{4^3}{x^3} & x > 4 \end{cases}$$

$$\begin{aligned} \mathrm{E}[X] &= \int_0^\infty S(x)dx = \int_0^4 1\,dx + \int_4^\infty \frac{4^3}{x^3}dx \\ &= 4 + \frac{-1}{2} \cdot \frac{4^3}{x^2}\Bigg|_4^\infty = 4 + \frac{4}{2} = \boxed{6} \end{aligned}$$

Or: X is a single parameter Pareto with $\alpha = 3$ and $\theta = 4$, so $\mathrm{E}[X] = \dfrac{\alpha\theta}{\alpha - 1} = \dfrac{3 \cdot 4}{3 - 1} = \boxed{6}$

- Outline

Moment Generating Functions (MGFs)

Rarely tested on Exam C, more of an Exam P topic.
On Exam C, it mostly comes up with frailty distributions.

Definition (MGF)

$M_X(t) = \mathrm{E}\left[e^{tX}\right] = \mathrm{E}\left[\left(e^t\right)^X\right]$

$$\begin{aligned} M_X(t) &= \mathrm{E}\left[e^{tX}\right] & M_X(0) &= \mathrm{E}\left[e^{0\cdot X}\right] = 1 \\ \frac{d}{dt}M_X(t) = M'_X(t) &= \mathrm{E}\left[Xe^{tX}\right] & M'_X(0) &= \mathrm{E}[X] \\ \frac{d^2}{dt^2}M_X(t) = M''_X(t) &= \mathrm{E}\left[X^2e^{tX}\right] & M''_X(0) &= \mathrm{E}\left[X^2\right] \\ &etc. \end{aligned}$$

Cumulant Generating Function

Even less likely than MGFs to be useful on Exam C.
Cumulant generating function give the 2nd and 3rd central moments.

$$\begin{aligned} g(t) &= \ln[M_X(t)] = \ln\left(\mathrm{E}\left[e^{tX}\right]\right) \\ g(0) &= \ln[M_X(0)] = \ln(1) = 0 \\ g'(0) &= \mu = \mathrm{E}[X] \\ g''(0) &= \sigma^2 = \mathrm{Var}[X] \\ g'''(0) &= \mathrm{E}[(X-\mu)^3] \\ g''''(0) &= \text{messy} \end{aligned}$$

This is rarely useful on Exam C. If we know the MGF, we typically know the distribution and therefore $\mathrm{Var}(X)$. When we want the 3rd central moment, we will rarely know the MGF so we will rarely (never?) use the cumulant generating function.

Probability Generating Functions

Tested more often (but still not every exam)

Definition (Probability Generating Function)

$P_X(z) = \mathrm{E}\left[z^X\right] = \mathrm{E}\left[\left(e^{\ln z}\right)^X\right] = M_X(\ln z)$

If X is discrete, then

$$\begin{aligned}
P_X(z) &= 1 \cdot \mathrm{P}[X=0] + z \cdot \mathrm{P}[X=1] + z^2 \cdot \mathrm{P}[X=2] + \dots \\
P(0) &= 1 \cdot \mathrm{P}[X=0] \\
P'(z) &= \quad 0 \quad + \quad \mathrm{P}[X=1] + 2z \cdot \mathrm{P}[X=2] + \dots \\
P'(0) &= \mathrm{P}[X=1] \\
P''(0) &= 2\mathrm{P}[X=2], \quad \mathrm{P}[X=2] = \frac{P''(0)}{2}
\end{aligned}$$

$$\mathrm{P}[X=k] = P_X^{(k)}(0)/k!$$

Exam tables include the MGF for some continuous distributions, and the PGF for discrete distributions.

PGFs and Moments

$M^{(k)}(0) = \mathrm{E}[X^k]$.

We can also find moments by taking derivatives of the PGF at 1.

$$\begin{aligned}
P_X(z) &= \mathrm{E}\left[z^X\right] & P_X(1) &= \mathrm{E}\left[1^X\right] = 1 \\
P'_X(z) &= \mathrm{E}\left[X z^{X-1}\right] & P'_X(1) &= \mathrm{E}[X] \\
P''_X(z) &= \mathrm{E}\left[X(X-1)z^{X-2}\right] & P''_X(1) &= \mathrm{E}[X(X-1)]
\end{aligned}$$

The derivatives at 1 are called "factorial" moments of X.

Example

Let N be a discrete random variable with

$$p_1 = 0.2 \quad p_2 = 0.4 \quad p_3 = 0.3 \quad p_4 = 0.1$$

What is the probability generating function of N? What is the second moment?

$$\begin{aligned}
P_N(z) &= \mathrm{E}[z^N] = 0.2 \cdot z^1 + 0.4 \cdot z^2 + 0.3 \cdot z^3 + 0.1 \cdot z^4 \\
\mathrm{E}[N^2] &= 0.2 \cdot 1^2 + 0.4 \cdot 2^2 + 0.3 \cdot 3^2 + 0.1 \cdot 4^2 = \boxed{6.1} \\
\text{Or: } P'(z) &= 0.2 + 0.8z + 0.9z^2 + 0.4z^3 \\
P''(z) &= \;0\; + 0.8 + 1.8z + 1.2z^2 \\
P'(1) &= 0.2 + 0.8 + 0.9 + 0.4 = 2.3 \\
P''(1) &= 0.8 + 1.8 + 1.2 = 3.8 \\
\mathrm{E}[N(N-1)] &= \mathrm{E}[N^2] - \mathrm{E}[N] \\
3.8 &= \mathrm{E}[N^2] - 2.3 \\
\mathrm{E}[N^2] &= \boxed{6.1}
\end{aligned}$$

Exercise 1

The probability generating function for a discrete variable N satisfies the following:

$$P'(0) = \frac{3}{16} \quad P'(1) = 3 \quad P''(0) = \frac{9}{32} \quad P''(1) = 18$$

Find $\mathrm{P}[N = 2]$.

Exercise 1

The probability generating function for a discrete variable N satisfies the following:

$$P'(0) = \frac{3}{16} \quad P'(1) = 3 \quad P''(0) = \frac{9}{32} \quad P''(1) = 18$$

Find $\mathrm{P}[N = 2]$.

$$\begin{aligned}
P_N(z) &= \mathrm{P}[N=0] + z \cdot \mathrm{P}[N=1] + z^2 \cdot \mathrm{P}[N=2] + z^3 \cdot \mathrm{P}[N=3] + \dots \\
P_N'(z) &= \quad 0 \quad + \mathrm{P}[N=1] + 2z \cdot \mathrm{P}[N=2] + 3z^2 \cdot \mathrm{P}[N=3] + \dots \\
P_N''(z) &= \quad 0 \quad + 2 \cdot \mathrm{P}[N=2] + 6z \cdot \mathrm{P}[N=3] + \dots \\
P_N''(0) &= 2 \cdot \mathrm{P}[N=2] \\
\frac{9}{32} &= 2 \cdot \mathrm{P}[N=2]
\end{aligned}$$

$$\mathrm{P}[N = 2] = \boxed{\frac{9}{64}}$$

Exercise 2

The probability generating function for a discrete distribution satisfies the following:

$$P'(0) = \frac{3}{16} \quad P'(1) = 3 \quad P''(0) = \frac{9}{32} \quad P''(1) = 18$$

Calculate the second moment of the distribution.

Exercise 2

The probability generating function for a discrete distribution satisfies the following:

$$P'(0) = \frac{3}{16} \quad P'(1) = 3 \quad P''(0) = \frac{9}{32} \quad P''(1) = 18$$

Calculate the second moment of the distribution.

$$\begin{aligned} \mathrm{E}[X] &= P'(1) = 3 \\ \mathrm{E}[X(X-1)] &= P''(1) = 18 \\ \mathrm{E}\left[X^2 - X\right] &= \mathrm{E}\left[X^2\right] - \mathrm{E}[X] = 18 \\ \mathrm{E}\left[X^2\right] &= 18 + \mathrm{E}[X] = \boxed{21} \end{aligned}$$

- Outline

Conditional Probability

$$\mathrm{P}[AB] = \mathrm{P}[A] \cdot \mathrm{P}[B \mid A]$$

$$\mathrm{P}[B \mid A] = \frac{\mathrm{P}[AB]}{\mathrm{P}[A]}$$

$$\mathrm{P}[A \mid B] = \frac{\mathrm{P}[AB]}{\mathrm{P}[B]}$$

If $A_1, A_2, \ldots, A_n$ are a *partition* of the sample space, meaning $A_i \cap A_j = \emptyset$ for all $i \neq j$ and $\sum \mathrm{P}[A_i] = 1$,

Theorem (Law of Total Probability)

$$P[B] = \sum_j P[A_j] \cdot P[B \mid A_j]$$

Theorem (Bayes' Theorem)

$$P[A_i \mid B] = \frac{P[A_i B]}{P[B]} = \frac{P[A_i] \cdot P[B \mid A_i]}{\sum_j P[A_j] \cdot P[B \mid A_j]}$$

Example

Low risk individuals have losses that are exponential with mean 100, while high risk individuals have losses that are exponential with mean 200. You observe a loss that is greater than 300. If 70% of losses are from low risk individuals, what is the probability that the observed loss was from a low risk individual?

Let X = loss amount, L denote low risk, and H high risk.

$$\begin{aligned}
\mathrm{P}[\mathrm{L} \mid X > 300] &= \frac{\mathrm{P}[\mathrm{L}, X > 300]}{\mathrm{P}[X > 300]} \\
&= \frac{\mathrm{P}[\mathrm{L}] \cdot \mathrm{P}[X > 300 \mid \mathrm{L}]}{\mathrm{P}[\mathrm{L}] \cdot \mathrm{P}[X > 300 \mid \mathrm{L}] + \mathrm{P}[\mathrm{H}] \cdot \mathrm{P}[X > 300 \mid \mathrm{H}]} \\
&= \frac{0.7 \cdot e^{-300/100}}{0.7 \cdot e^{-300/100} + 0.3 \cdot e^{-300/200}} \\
&= \boxed{0.342}
\end{aligned}$$

Discrete Example

The joint distribution of X and Y is given by

		X		
		0	1	2
Y	1	0.1	0.2	0.3
	2	0.1	0.1	0.2

$$\mathrm{P}[X=0] = 0.1 + 0.1 = 0.2$$

$$\mathrm{P}[X=1] = 0.2 + 0.1 = 0.3$$

$$\mathrm{P}[X=2] = 0.3 + 0.2 = 0.5$$

$$\mathrm{P}[Y=1] = 0.1 + 0.2 + 0.3 = 0.6$$

$$\mathrm{P}[Y=2] = 0.1 + 0.1 + 0.2 = 0.4$$

$$\mathrm{P}[X=0 \mid Y=1] = \frac{0.1}{0.6} = \frac{1}{6}$$

$$\mathrm{P}[X=1 \mid Y=1] = \frac{0.2}{0.6} = \frac{2}{6}$$

$$\mathrm{P}[X=2 \mid Y=1] = \frac{0.3}{0.6} = \frac{3}{6}$$

$$\mathrm{P}[Y=1 \mid X=2] = 0.6$$

$$\mathrm{P}[Y=2 \mid X=2] = 0.4$$

Formulas

Discrete: $$\mathrm{P}[X=x] = \sum_y \mathrm{P}[X=x, Y=y]$$

$$\mathrm{P}[X=x \mid Y=y] = \frac{\mathrm{P}[X=x, Y=y]}{\mathrm{P}[Y=y]} = \frac{\mathrm{P}[X=x, Y=y]}{\sum_x \mathrm{P}[X=x, Y=y]}$$

Continuous: $$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy$$

$$f_{X|Y}(x \mid Y=y) = \frac{f(x,y)}{f_Y(y)} = \frac{f(x,y)}{\int_{-\infty}^{\infty} f(x,y)\,dx}$$

X and Y are independent if and only if 1) $f(x,y) = f(x) \cdot f(y)$ and 2) the range of X and Y is rectangular.

Exercise 1

A blood test indicates the presence of a particular disease 90% of the time when the disease is actually present. The same test indicates the presence of the disease 2% of the time when the disease is not present. One percent of the population actually has the disease. Calculate the probability that a person has the disease given that the test indicates the presence of the disease.

Exercise 1

A blood test indicates the presence of a particular disease 90% of the time when the disease is actually present. The same test indicates the presence of the disease 2% of the time when the disease is not present. One percent of the population actually has the disease. Calculate the probability that a person has the disease given that the test indicates the presence of the disease.

Let T denote a positive test, and D the disease.

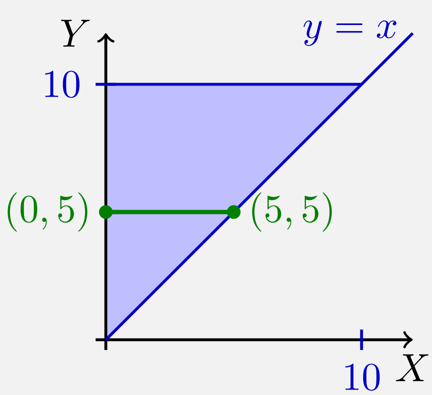
$$
\begin{aligned}
\mathrm{P}[D \mid T] &= \frac{\mathrm{P}[DT]}{\mathrm{P}[T]} \\
&= \frac{\mathrm{P}[D] \cdot \mathrm{P}[T \mid D]}{\mathrm{P}[D] \cdot \mathrm{P}[T \mid D] + \mathrm{P}[D'] \cdot \mathrm{P}[T \mid D']} \\
&= \frac{0.01 \cdot 0.90}{0.01 \cdot 0.90 + 0.99 \cdot 0.02} \\
&= \boxed{0.3125}
\end{aligned}
$$

Exercise 2

X and Y have joint density $0.3x^2y^{-3}$ for $0 < x < y < 10$. Find $\mathrm{P}[X \le 3 \mid Y = 5]$.

Exercise 2

X and Y have joint density $0.3x^2y^{-3}$ for $0 < x < y < 10$. Find $\mathrm{P}[X \le 3 \mid Y = 5]$.



$$f_{X|Y}(x \mid Y = 5) = \frac{f(x,5)}{\int_0^5 f(x,5)dx} = \frac{0.3x^2 5^{-3}}{\int_0^5 0.3x^2 5^{-3} dx} = \frac{x^2}{\int_0^5 x^2\,dx} = \frac{3x^2}{125}$$

$$\mathrm{P}[X \le 3 \mid Y = 5] = \int_0^3 \frac{3x^2}{125} dx = \boxed{\frac{27}{125}}$$

Questions like this are very rare on the exam.

- Outline

Law of Total Expectation

For a discrete variable, $\mathrm{E}[g(X)] = \sum g(x) \cdot \mathrm{P}[X = x]$

If $A_1, \ldots, A_n$ are a partition of the sample space, then for any variable

$$\mathrm{E}[X] = \sum_i \mathrm{E}[X \mid A_i] \cdot \mathrm{P}[A_i]$$
$$\mathrm{E}[X^k] = \sum_i \mathrm{E}[X^k \mid A_i] \cdot \mathrm{P}[A_i]$$

Example: Low risk individuals have average annual losses of 100, high risk have average annual losses of 300.
If 40% of a group are low risk and 60% high risk, then the average for the group is $0.4 \cdot 100 + 0.6 \cdot 300 = 220$

MultiVariate Example

		X		
		0	1	2
Y	1	0.1	0.2	0.3
	2	0.1	0.1	0.2

$$\begin{aligned} \mathrm{E}[X] &= 0 \cdot 0.1 + 1 \cdot 0.2 + 2 \cdot 0.3 + 0 \cdot 0.1 + 1 \cdot 0.1 + 2 \cdot 0.2 \\ &= 0 \cdot 0.2 + 1 \cdot 0.3 + 2 \cdot 0.5 \\ &= 0 \cdot \mathrm{P}[X = 0] + 1 \cdot \mathrm{P}[X = 1] + 2 \cdot \mathrm{P}[X = 2] \\ &= 1.3 \end{aligned}$$

Last time, we saw

$$\mathrm{P}[X = 0 \mid Y = 1] = \frac{1}{6} \quad \mathrm{P}[X = 1 \mid Y = 1] = \frac{2}{6} \quad \mathrm{P}[X = 2 \mid Y = 1] = \frac{3}{6}$$

$$\mathrm{E}[X \mid Y = 1] = 0 \cdot \frac{1}{6} + 1 \cdot \frac{2}{6} + 2 \cdot \frac{3}{6} = \frac{4}{3}$$

Likewise, $\mathrm{E}[X \mid Y = 2] = \frac{5}{4}$

Example Continued

$$\mathrm{E}[X \mid Y = 1] = \frac{4}{3} \qquad \mathrm{E}[X \mid Y = 2] = \frac{5}{4}$$

Key point: $\mathrm{E}[X \mid Y]$ is a function of Y. *As a result, it is also a random variable.*

$$\begin{aligned} \mathrm{P}[\mathrm{E}[X \mid Y] = 4/3] &= \mathrm{P}[Y = 1] = 0.6 \\ \mathrm{P}[\mathrm{E}[X \mid Y] = 5/4] &= \mathrm{P}[Y = 2] = 0.4 \\ \mathrm{E}[\mathrm{E}[X \mid Y]] &= 0.6 \cdot \frac{4}{3} + 0.4 \cdot \frac{5}{4} \\ &= 1.3 \\ &= \mathrm{E}[X] \end{aligned}$$

That is not a coincidence.
Remember: $\mathrm{E}[X \mid Y]$ is random. $\mathrm{E}[X]$ is a (non-random) number.

Double Expectation

Recall that

$$E[g(Y)] = \sum_y g(y) \cdot P[Y = y]$$

$$E[X] = \sum_i E[X \mid A_i] \cdot P[A_i]$$

$$E[X] = \sum_y E[X \mid Y = y] \cdot P[Y = y]$$

$$= E[E[X \mid Y]]$$

This works for all raw moments:

$$E[X^k] = E[E[X^k \mid Y]]$$

Central moments require an adjustment:

$$\mathrm{Var}[X] = E[\mathrm{Var}[X \mid Y]] + \mathrm{Var}[E[X \mid Y]]$$

Example

		X		
		0	1	2
Y	1	0.1	0.2	0.3
	2	0.1	0.1	0.2

Find Var[X]

$$E[X \mid Y = 1] = \frac{4}{3} \quad E[X \mid Y = 2] = \frac{5}{4}$$

$$\mathrm{Var}[X \mid Y = 1] = 0^2 \cdot \frac{1}{6} + 1^2 \cdot \frac{2}{6} + 2^2 \cdot \frac{3}{6} - \left(\frac{4}{3}\right)^2 = \frac{5}{9}$$

$$\mathrm{Var}[X \mid Y = 2] = 0^2 \cdot \frac{1}{4} + 1^2 \cdot \frac{1}{4} + 2^2 \cdot \frac{2}{4} - \left(\frac{5}{4}\right)^2 = \frac{11}{16}$$

$$\mathrm{Var}[X] = E[\mathrm{Var}[X \mid Y]] + \mathrm{Var}[E[X \mid Y]]$$

$$= 0.6 \cdot \frac{5}{9} + 0.4 \cdot \frac{11}{16} + \left(\frac{4}{3} - \frac{5}{4}\right)^2 \cdot 0.6 \cdot 0.4$$

$$= \boxed{0.61}$$

Exercise 1

I roll a fair six-sided die, and N comes up. I then flip N independent fair coins. Find the mean and variance of the number of heads.

Exercise 1

I roll a fair six-sided die, and N comes up. I then flip N independent fair coins. Find the mean and variance of the number of heads.

Let $H = \#$of heads. If we knew N this would be easy. That is a hint to use double expectation.

$$\begin{aligned}
\mathrm{E}[H] &= \mathrm{E}[\mathrm{E}[H \mid N]] \\
&= \mathrm{E}\left[\frac{N}{2}\right] = \frac{1}{2} \cdot \mathrm{E}[N] \\
&= \frac{1}{2} \cdot \frac{1+6}{2} = \boxed{\frac{7}{4}} \\
\mathrm{Var}[H] &= \mathrm{E}[\mathrm{Var}[H \mid N]] + \mathrm{Var}[\mathrm{E}[H \mid N]] \\
&= \mathrm{E}\left[N \cdot \frac{1}{2} \cdot \frac{1}{2}\right] + \mathrm{Var}\left[\frac{N}{2}\right] \\
&= \frac{1}{4}\mathrm{E}[N] + \frac{1}{4}\mathrm{Var}[N] = \frac{1}{4} \cdot \frac{7}{2} + \frac{1}{4} \cdot \frac{35}{12} = \boxed{\frac{77}{48}}
\end{aligned}$$

Exercise 2

The joint distribution of X and Y is as follows. Find $\mathrm{E}[Y]$ and $\mathrm{Var}[Y]$.

		X		
		0	1	2
Y	1	0.1	0.2	0.3
	2	0.1	0.1	0.2

Exercise 2

The joint distribution of X and Y is as follows. Find $\mathrm{E}[Y]$ and $\mathrm{Var}[Y]$.

		X		
		0	1	2
Y	1	0.1	0.2	0.3
	2	0.1	0.1	0.2

$$\begin{aligned}
\mathrm{E}[Y] &= 1 \cdot (0.1 + 0.2 + 0.3) + 2 \cdot (0.1 + 0.1 + 0.2) \\
&= 1 \cdot 0.6 + 2 \cdot 0.4 = \boxed{1.4} \\
\mathrm{E}[Y^2] &= 1^2 \cdot 0.6 + 2^2 \cdot 0.4 = 2.2 \\
\mathrm{Var}[Y] &= 2.2 - 1.4^2 = \boxed{0.24}
\end{aligned}$$

$$\text{Or: } \mathrm{E}[Y \mid X = 0] = 1 \cdot \frac{0.1}{0.1 + 0.1} + 2 \cdot \frac{0.1}{0.1 + 0.1} = 1.5$$

$$\mathrm{E}[Y \mid X = 1] = 1 \cdot \frac{0.2}{0.2 + 0.1} + 2 \cdot \frac{0.1}{0.2 + 0.1} = \frac{4}{3}$$

$$\mathrm{E}[Y \mid X = 2] = 1 \cdot \frac{0.3}{0.3 + 0.2} + 2 \cdot \frac{0.2}{0.3 + 0.2} = 1.4$$

$$\mathrm{E}[Y] = \mathrm{E}[\mathrm{E}[Y \mid X]] = 1.5 \cdot 0.2 + \frac{4}{3} \cdot 0.3 + 1.4 \cdot 0.5 = \boxed{1.4}$$

Exercise 2 (Continued)

		X		
		0	1	2
Y	1	0.1	0.2	0.3
	2	0.1	0.1	0.2

$$\mathrm{E}[Y \mid X = 0] = 1.5$$
$$\mathrm{E}[Y \mid X = 1] = 4/3$$
$$\mathrm{E}[Y \mid X = 2] = 1.4$$

$$\mathrm{Var}[Y \mid X = 0] = (2-1)^2 \cdot 0.5 \cdot 0.5 = 0.25$$
$$\mathrm{Var}[Y \mid X = 1] = (2-1)^2 \cdot \frac{2}{3} \cdot \frac{1}{3} = \frac{2}{9}$$
$$\mathrm{Var}[Y \mid X = 2] = (2-1)^2 \cdot 0.6 \cdot 0.4 = 0.24$$

$$\begin{aligned}\mathrm{Var}[Y] &= \mathrm{E}[\mathrm{Var}[Y \mid X]] + \mathrm{Var}[\mathrm{E}[Y \mid X]] \\ &= 0.2 \cdot 0.25 + 0.3 \cdot \frac{2}{9} + 0.5 \cdot 0.24 \\ &\quad + \left(1.5^2 \cdot 0.2 + \frac{4^2}{3^2} \cdot 0.3 + 1.4^2 \cdot 0.5\right) - 1.4^2 \\ &= \boxed{0.24}\end{aligned}$$

- Outline

Example

An insurance company has two types of customers: high risk customers, whose annual losses are exponential with mean 10, and low risk customers, whose annual losses are exponential with mean 1. 30% of the customers are high risk.
A customer is chosen at random. What is the probability that their annual loss is at least 5?

Let L denote low risk, H high risk, X annual losses.

$$\begin{aligned} \mathrm{P}[X \geq 5] &= \mathrm{P}[L, X \geq 5] + \mathrm{P}[H, X \geq 5] \\ &= \mathrm{P}[L] \cdot \mathrm{P}[X \geq 5 \mid L] + \mathrm{P}[H] \cdot \mathrm{P}[X \geq 5 \mid H] \\ &= 0.7 \cdot e^{-5/1} + 0.3 \cdot e^{-5/10} \\ &= \boxed{0.187} \end{aligned}$$

n-point Mixture

Suppose we have n cases.
Exactly 1 case occurs.
Case i: $Y = X_i$, $\mathrm{P}[\text{case } i] = a_i$, $\sum a_i = 1$

$$\begin{aligned} F_Y(y) &= a_1 F_{X_1}(y) + \cdots + a_n F_{X_n}(y) \\ f_Y(y) &= a_1 f_{X_1}(y) + \cdots + a_n f_{X_n}(y) \\ \mathrm{E}[Y] &= \mathrm{E}[\mathrm{E}[Y \mid \text{case}]] = a_1 \mathrm{E}[X_1] + \cdots + a_n \mathrm{E}[X_n] \\ \mathrm{E}[Y^k] &= a_1 \mathrm{E}[X_1^k] + \cdots + a_n \mathrm{E}[X_n^k] \\ \mathrm{Var}[Y] &= \mathrm{E}[\mathrm{Var}[Y \mid \text{Case}]] + \mathrm{Var}[\mathrm{E}[Y \mid \text{Case}]] \end{aligned}$$

Warning: $Y \neq \sum a_i X_i$
We will discuss the differences next lesson.

Example

An insurance company has two types of customers: high risk customers, whose annual losses have mean 10 and variance 50, and low risk customers, whose annual losses have mean 5 and variance 20. 30% of the customers are high risk.
Find the mean and variance of the annual loss of a randomly chosen customer.

Let Y be a random annual loss, L = low risk, H = high risk.

$$\begin{aligned}
\mathrm{E}[Y] &= \mathrm{E}[\mathrm{E}[Y \mid \text{Case}]] \\
&= \mathrm{P}[H] \cdot \mathrm{E}[Y \mid H] + \mathrm{P}[L] \cdot \mathrm{E}[Y \mid L] \\
&= 0.3 \cdot 10 + 0.7 \cdot 5 = \boxed{6.5} \\
\mathrm{Var}[Y] &= \mathrm{E}[\mathrm{Var}[Y \mid \text{Case}]] + \mathrm{Var}[\mathrm{E}[Y \mid \text{Case}]] \\
&= 0.3 \cdot 50 + 0.7 \cdot 20 + (10 - 5)^2 \cdot 0.3 \cdot 0.7 \\
&= \boxed{34.25}
\end{aligned}$$

Example

An insurance company has 3 types of customers, whose annual losses are described below. Find the mean and variance of the annual loss of a randomly chosen customer.

Type	Proportion of Customers	Average Annual Loss	Variance of Annual Losses
Low	60%	5	20
Med	30%	10	50
High	10%	30	100

Y = random annual loss, L = low risk, M = med. risk, H = high risk.

$$\begin{aligned}
\mathrm{E}[Y] &= 0.6 \cdot \mathrm{E}[Y \mid L] + 0.3 \cdot \mathrm{E}[Y \mid M] + 0.1 \cdot \mathrm{E}[Y \mid H] \\
&= 0.6 \cdot 5 + 0.3 \cdot 10 + 0.1 \cdot 30 = \boxed{9} \\
\mathrm{E}[\mathrm{Var}[Y \mid \text{Case}]] &= 0.6 \cdot 20 + 0.3 \cdot 50 + 0.1 \cdot 100 = 37 \\
\mathrm{Var}[\mathrm{E}[Y \mid \text{Case}]] &= 5^2 \cdot 0.6 + 10^2 \cdot 0.3 + 30^2 \cdot 0.1 \\
&\quad - (5 \cdot 0.6 + 10 \cdot 0.3 + 30 \cdot 0.1)^2 \\
&= 135 - 9^2 = 54 \\
\mathrm{Var}[Y] &= 37 + 54 = \boxed{91}
\end{aligned}$$

Exercise 1

An insurance company has two types of customers: high risk customers, whose loss amounts have mean 10 and variance 50, and low risk customers, whose losses have mean 5 and variance 20. The expected amount of a randomly chosen loss is 8.
Find the variance of a randomly chosen loss.

Exercise 1

An insurance company has two types of customers: high risk customers, whose loss amounts have mean 10 and variance 50, and low risk customers, whose losses have mean 5 and variance 20. The expected amount of a randomly chosen loss is 8.
Find the variance of a randomly chosen loss.

Y = random loss, L = low risk, H = high risk. Let $p = \mathrm{P}[L]$.

$$\begin{aligned} \mathrm{E}[Y] &= 8 = p \cdot 5 + (1-p) \cdot 10 \\ p &= 0.4 \\ \mathrm{Var}[Y] &= \mathrm{E}[\mathrm{Var}[Y \mid \mathrm{Case}]] + \mathrm{Var}[\mathrm{E}[Y \mid \mathrm{Case}]] \\ &= 0.4 \cdot 20 + 0.6 \cdot 50 + (10-5)^2 \cdot 0.4 \cdot 0.6 \\ &= \boxed{44} \end{aligned}$$

Exercise 2

Rural losses have hazard rate $h(x) = 0.2x$ for $x > 0$, while urban losses have hazard rate $h(x) = 0.03x^2$ for $x > 0$. If 30% of losses are rural, what proportion of losses are at most 5?

Exercise 2

Rural losses have hazard rate $h(x) = 0.2x$ for $x > 0$, while urban losses have hazard rate $h(x) = 0.03x^2$ for $x > 0$. If 30% of losses are rural, what proportion of losses are at most 5?
Let R denote rural, U urban, and Y a random loss.

$$\begin{aligned}
\mathrm{P}[Y \le 5] &= \mathrm{P}[R] \cdot \mathrm{P}[Y \le 5 \mid R] + \mathrm{P}[U] \cdot \mathrm{P}[Y \le 5 \mid U] \\
&= 0.3 \cdot \mathrm{P}[Y \le 5 \mid R] + 0.7 \cdot \mathrm{P}[Y \le 5 \mid U] \\
\mathrm{P}[Y \le 5 \mid R] &= 1 - \mathrm{P}[Y > 5 \mid R] \\
&= 1 - e^{-\int_0^5 0.2x\,dx} = 1 - e^{-0.1 \cdot 5^2} = 0.918 \\
\mathrm{P}[Y \le 5 \mid U] &= 1 - e^{-\int_0^5 0.03x^2\,dx} \\
&= 1 - e^{-0.01 \cdot 5^3} = 0.713 \\
\mathrm{P}[Y \le 5] &= 0.3 \cdot 0.918 + 0.7 \cdot 0.713 = \boxed{0.775}
\end{aligned}$$

- Outline

Sums of Independent Variables

For any random variables,

$$\mathrm{E}[X+Y] = \mathrm{E}[X] + \mathrm{E}[Y]$$
$$\mathrm{E}\left[\sum X_i\right] = \sum \mathrm{E}[X_i]$$

If in addition the variables are independent then

$$\mathrm{Var}[X+Y] = \mathrm{Var}[X] + \mathrm{Var}[Y]$$
$$\mathrm{Var}\left[\sum X_i\right] = \sum \mathrm{Var}[X_i]$$

What if X and Y are not independent?

Covariance

$$
\begin{aligned}
\mathrm{Var}[X] &= \mathrm{E}[(X-\mu)^2] = \mathrm{E}[X^2] - (\mathrm{E}[X])^2 \\
\mathrm{Cov}[X,Y] &= \mathrm{E}[(X-\mu_X)(Y-\mu_Y)] = \mathrm{E}[XY] - \mathrm{E}[X]\cdot\mathrm{E}[Y] \\
\mathrm{Var}[X+Y] &= \mathrm{E}[(X+Y)^2] - (\mathrm{E}[X+Y])^2 \\
&= \mathrm{E}[X^2+2XY+Y^2] - \big((\mathrm{E}[X])^2 + 2\mathrm{E}[X]\mathrm{E}[Y] + \mathrm{E}[Y])^2\big) \\
&= \mathrm{E}[X^2] - (\mathrm{E}[X])^2 + 2(\mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y]) \\
&\qquad + \mathrm{E}[Y^2] - (\mathrm{E}[Y])^2 \\
&= \mathrm{Var}[X] + 2\,\mathrm{Cov}[X,Y] + \mathrm{Var}[Y]
\end{aligned}
$$

If X and Y are independent, then $\mathrm{E}[XY] = \mathrm{E}[X]\mathrm{E}[Y]$ and $\mathrm{Cov}[X,Y] = 0$. So if X and Y are independent

$\mathrm{Var}[X+Y] = \mathrm{Var}[X] + \mathrm{Var}[Y]$

Properties of Covariance

Covariance is a “bilinear form” meaning that it follows the usual distributive laws.

$$
\begin{aligned}
(aX+bY)(cZ+dW) &= acXZ + adXW + bcYZ + bdYW \\
\mathrm{Cov}(aX+bY, cZ+dW) &= ac\,\mathrm{Cov}(X,Z) + ad\mathrm{Cov}(X,W) \\
&\qquad + bc\,\mathrm{Cov}(Y,Z) + bd\mathrm{Cov}(Y,W) \\
\mathrm{Cov}(X,X) &= \mathrm{E}[X\cdot X] - \mathrm{E}[X]\cdot\mathrm{E}[X] = \mathrm{Var}[X] \\
(aX+bY)^2 &= a^2X^2 + 2abXY + b^2Y^2 \\
\mathrm{Var}[aX+bY] &= a^2\mathrm{Var}[X] + 2ab\,\mathrm{Cov}(X,Y) + b^2\mathrm{Var}[Y]
\end{aligned}
$$

Recall the correlation of X and Y is

$$
\begin{aligned}
\mathrm{Corr}(X,Y) &= \frac{\mathrm{Cov}(X,Y)}{\mathrm{SD}[X]\mathrm{SD}[Y]} \\
-1 \le \mathrm{Corr}(X,Y) &\le 1
\end{aligned}
$$

Mixtures vs Sums

Example

A homeowners policy covers flood and theft losses. The amount of each flood loss is exponential with mean 10, the amount of each theft loss is exponential with mean 20, and a homeowner suffers one flood loss and one theft loss.
If the loss amounts are independent, and the insurance payment is 40% of the flood loss and 60% of the theft loss, what is the mean and variance of the payment?

Let $F \sim \exp(10)$ = flood losses, and $T \sim \exp(20)$ = theft losses. The payment is $0.4F + 0.6T$, and

$$\mathrm{E}[0.4F + 0.6T] = 0.4 \cdot 10 + 0.6 \cdot 20 = 16$$
$$\mathrm{Var}[0.4F + 0.6T] = 0.4^2 \cdot 10^2 + 0.6^2 \cdot 20^2 = 160$$

This was a sum of two random variables, not a mixture.

Example

A homeowners policy covers flood and theft losses. The amount of each flood loss is exponential with mean 10 and the amount of each theft loss is exponential with mean 20. If 40% of losses are flood losses, and 60% are theft losses, what is the mean and variance of a randomly selected loss amount?

This is a mixture because a randomly selected loss is either a flood loss, or a theft loss, but not both at the same time.
Let X denote the loss amount.

Mixture example

With probability 0.4, $X \sim \exp(10)$,
and with probability 0.6, $X \sim \exp(20)$

$$\begin{aligned}
\mathrm{E}[X] &= \mathrm{E}[\mathrm{E}[X \mid \text{Case}]] \\
&= 0.4 \cdot 10 + 0.6 \cdot 20 = 16 \\
\mathrm{Var}[X] &= \mathrm{E}[\mathrm{Var}(X \mid \text{Case})] + \mathrm{Var}[\mathrm{E}(X \mid \text{Case})] \\
&= 0.4 \cdot 100 + 0.6 \cdot 400 + (20-10)^2 \cdot 0.4 \cdot 0.6 \\
&= 304 \\
\mathrm{P}[X \leq x] &= 0.4\mathrm{P}[F \leq x] + 0.6\mathrm{P}[T \leq x] \\
&= 0.4\left(1 - e^{-x/10}\right) + 0.6\left(1 - e^{-x/20}\right) \\
&= 1 - 0.4e^{-x/10} - 0.6e^{-x/20}
\end{aligned}$$

Exercise 1

$\mathrm{E}[X] = 2, \mathrm{E}[Y] = 3, \mathrm{Var}[X] = 1, \mathrm{Var}[Y] = 5, \mathrm{Var}[X + 2Y] = 13.$
Find $\mathrm{E}[XY]$.

Exercise 1

$\mathrm{E}[X] = 2, \mathrm{E}[Y] = 3, \mathrm{Var}[X] = 1, \mathrm{Var}[Y] = 5, \mathrm{Var}[X + 2Y] = 13.$
Find $\mathrm{E}[XY]$.

$$\begin{aligned} \mathrm{Var}[X + 2Y] &= \mathrm{Var}[X] + 2\mathrm{Cov}(X, 2Y) + \mathrm{Var}[2Y] \\ &= \mathrm{Var}[X] + 4\mathrm{Cov}(X, Y) + 4\mathrm{Var}[Y] \\ 13 &= 1 + 4 \cdot \mathrm{Cov}(X, Y) + 4 \cdot 5 \\ \mathrm{Cov}(X, Y) &= -2 \\ \mathrm{Cov}(X, Y) &= \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y] \\ -2 &= \mathrm{E}[XY] - 2 \cdot 3 \\ \boxed{4} &= \mathrm{E}[XY] \end{aligned}$$

Exercise 2

Let X and Y be independent variables with $\mathrm{E}[X] = 3$, $\mathrm{E}[Y] = 5$, $\mathrm{Var}[X] = \mathrm{Var}[Y] = 4$. Suppose that $W = 0.4X + 0.6Y$, and let Z be a variable that equals X 40% of the time and Y 60% of the time. Find $\mathrm{E}[W], \mathrm{E}[Z], \mathrm{Var}[W]$ and $\mathrm{Var}[Z]$.

Exercise 2

Let X and Y be independent variables with $\mathrm{E}[X] = 3$, $\mathrm{E}[Y] = 5$, $\mathrm{Var}[X] = \mathrm{Var}[Y] = 4$. Suppose that $W = 0.4X + 0.6Y$, and let Z be a variable that equals X 40% of the time and Y 60% of the time. Find $\mathrm{E}[W], \mathrm{E}[Z], \mathrm{Var}[W]$ and $\mathrm{Var}[Z]$.

Remark: W is a sum, Z is a mixture.

$$
\begin{aligned}
\mathrm{E}[W] &= \mathrm{E}[0.4X + 0.6Y] = 0.4\mathrm{E}[X] + 0.6\mathrm{E}[Y] = 4.2 \\
\mathrm{E}[Z] &= 0.4\mathrm{E}[X] + 0.6\mathrm{E}[Y] = 4.2 \\
\mathrm{Var}[W] &= 0.4^2\mathrm{Var}[X] + 2 \cdot 0.4 \cdot 0.6\mathrm{Cov}(X, Y) + 0.6^2\mathrm{Var}[Y] \\
&= 0.16 \cdot 4 + 0 + 0.36 \cdot 4 = 2.08 \\
\mathrm{Var}[Z] &= \mathrm{E}[\mathrm{Var}[Z \mid \mathrm{Case}]] + \mathrm{Var}[\mathrm{E}[Z \mid \mathrm{Case}]] \\
&= 0.4 \cdot 4 + 0.6 \cdot 4 + (5 - 3)^2 \cdot 0.4 \cdot 0.6 \\
&= 4.96
\end{aligned}
$$