§3.1 RECALL: DEFINITION OF RANDOM VARIABLE:

MRT2 §5.1	A variable (X) whose value is a number determined by the outcome of an experiment Can also be considered as a *function* that assigns a real number to each sample point i.e. {X(e_1), X(e_2), ... }
MM3 §4.3	A variable (X) whose value is a numerical outcome of a random phenomenon
WMS5 §2.11	A real-valued function for which the domain is a sample space. [Y: variable to be measured]

§3.1 DISCRETE RANDOM VARIABLE

Text	Definition
WMS5 §3.1	A random variable that assumes only a finite (or countably infinite) number of distinct values
MRT2 §7.2	Discrete random variables have "a finite or countably infinite number of possible values, each with positive or zero probability.
MM3 §4.3	A discrete random variable X has a finite number of possible values

RANDOM VARIABLE: DISCRETE or CONTINUOUS?

How long you have to wait for bus / elevator / surgery/ download to complete; how many tries before pass a course; length of song on a CD ; how long to download a specific song from Napster; how hot it is going to be today; how much snow we will get next winter; time someone called (on answering machine); how much ice McDonalds puts in soft drink / calories in hamburger / how many correct in 6/49? where roulette wheel stops; how many "wrong number" calls received; how many keys have to try before get the right one; water consumed by 100 homes.

§3.2 PROBABILITY DISTRIBUTION p(y) ASSOCIATED WITH RANDOM VARIABLE (Y):

The ordered pairs {y, Probability(Y = y) } , where y ranges over the possible values of Y

- Probability(Y = y) often shortened to Prob(Y = y) or P(Y = y)
- Can display distribution as formula, table, graph, etc..

EXAMPLES

Put 3 events in the order in which they occurred	Y = Number in correct Position, if Guess y Probability(Y = y) 0 2/6 1 3/6 3 1/6
2 of 4 cans filled with water (W) Guess which 2 contain water	Y = Number of correctly identified Cans y Probability(Y = y) 0 1/6 1 4/6 2 1/6
Chose a word and measure how long it is, i.e., # characters (c's)	Y = Number of characters in word y Probability(Y = y) 1 p(1) 2 p(2) ... ...

It helps if one thinks of probability as a proportion

MORE PROBABILITY DISTRIBUTIONS

Experiment	R.V. and associated distribution

First Ace
 	
Its location in a well shuffled deck
	Y = location (from top)
	`y Probability(Y = y)`
	`(1-49)`
	`see spreadsheet`

When First Duplicate Birthday if ask people to tell their birthdays	Y = when get 1st duplicate *
	`y Probability(Y = y)`
	`(2-365)`
	`see spreadsheet`
	
	*If birthdays uniformly distributed over 365 days of year

la Quotidienne 3
 	
(same distribution for exact order and any order)
	Y = Winnings on a $1 wager *
	`y Probability(Y = y)`
	`$0 0.999`
	`$450 0.001`

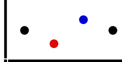
la Quotidienne 4
 	
(same distribution for exact order and any order)
	Y = Winnings on a $1 wager *
	`y Probability(Y = y)`
	`$0 0.9999`
	`$4500 0.0001`

Keno	See class web page

Winning margin in a game in **NHL** regular season	Y = Number of goals game won by
	`y Probability(Y = y)`
	`0 goals p(0)`
	`1 goal p(1)`
	`... ...`
	`etc.`

YET MORE PROBABILITY DISTRIBUTIONS

Experiment	R.V. and associated distribution

(**Turning Points** in Economic Time Series, MRT2)
 	
A value is a *turning point* if it is the least or greatest of itself and its immediate neighbours. e.g.
 (2 turns)
	If all permutations of 4 different measurements are equally likely, what is the probability distribution of the number of turning points?
	
	Y = number of turning points
	`y Probability(Y = y)`
	`0 2/24`
	`1 12/12`
	`2 10/24`

6/49 lottery
 	
Player selects 6 distinct numbers on a grid showing the numbers 1 to 49
 	
49 otherwise identical balls, but numbered 1 to 49, are thoroughly mixed in an urn; 6 balls are drawn without replacement.
	Y = how many of the balls drawn show numbers that match the numbers selected by player
	`y Probability(Y = y)`
	`0 0.4359650`
	`1 0.4130195`
	`2 0.1323780`
	`3 0.0176504`
	`4 0.0009686`
	`5 0.0000184`
	`6 0.0000001`

la Quotidienne 3
 	
(same distribution for exact order and any order)
	Y = Winnings on a $1 wager *
	`y Probability(Y = y)`
	`$0 0.999`
	`$450 0.001`

Keno	See class web page

EXPECTED VALUE OF DISCRETE RANDOM VARIABLE

See DEFINITION 3.4 (WMS5)

Y ~ p(y)

i.e., the r.v. Y has prob. distribution p(y)

$E(Y) = \Sigma_y \; y \; p(y)$,

i.e. weighted ave. of y's, with p(y)'s as weights (remember calculating mean length of words; here divisor is 1, [$\Sigma_y \; p(y) = 1$], rather than n.

Can think of E(Y) as "center of mass" of p()

la Quotidienne 3 some may find it easier to think of averaging the winnings of 1 who won $1 and 999 who won $0, or 1000 & 999999 …	Y = Winnings on a $1 wager * y p(y) y p(y) $0 0.999 $0.00 $450 0.001 $0.45 ----- Σ $0.45 i.e E(Winnings)=$0.45
Tri-State Pick 3 [see MM3, p 338) same as above, but payoff is $500 for exact order, $83.33 for any order (all ∂iff)	Y = Winnings on a $1 wager * E(Winnings)=$0.50
•Keno see class web page	•Y = Winnings on a $3 wager * E(Winnings) = $2.12 (> 70%)
•Banco see web page	Y = Winnings on a $1 wager * 55% ≤ E(Winnings) ≤ 55%

•Video lottery Machines	<<15 200 machines ont généré, en 1999-2000, des profits nets de 553 millions, soit 42% de la marge bénéficiare dégagée par le jeu sous toutes ses formes>> Mario Roy, La Presse (from The Gazette May 6, 2001)
First Ace see web page	Y = Location of First Ace E(Location) = 9.6
First Duplicate Birthday see web page	Y = when get 1st duplicate E(when) = 24.6 (approx)
Longevity of a fictitious new birth cohort if they were to experience age-specific death rates observed in Québec in1990 cf web page & exer.	Y = Age at death (a.k.a. Longevity or length of Life) E(Life) = 7x.x (approx) "LIFE EXPECTANCY AT BIRTH"
If you want to start your own Insurance Company see MM3 p 341	Y = Payout (from -99,750 to +$1250) E(Payout on single policy) > 0 BUT... Variance(Payout) is VERY LARGE
Where to wait if 3 unequally spaced elevators ? p(it's #1) = p(it's #2)=p(it's #3) = 1/3	Y = Distance to Elevator How to mimimize E(distance)? Cf my notes on MMCh 4

EXPECTED VALUE OF A FUNCTION OF A DISCRETE RANDOM VARIABLE

See THEOREM 3.2 (WMS5)

Y ~ p(y)

i.e., the r.v. Y has prob. distribution p(y)

g(Y) is some real-valued function of Y

E[**g**(Y)] $= \Sigma_y$ g(y) p(y),

i.e. weighted ave. of g(y)'s,
with p(y)'s as weights

EXAMPLES

Random Variable, Y	g(Y)
Temperature (C) on May 6 in Mtl	Temperature (F) = 32 + (9/5) Y
Weight in Kg <----------> Height in cm <---------->	Weight in Kg Height in inches
diameter of randomly chosen sphere	Volume of sphere = $(\pi/6)\ Y^3$
Y = CDN$1 in US$	US$1 in $CDN = 1/Y (Important that Y > 0 !!)
random Variable Y with Expectation μ ("mu")	$\lvert Y - \mu \rvert$, absolute deviation from the mean
random Variable Y with Expectation μ ("mu")	$(Y - \mu)^2$, squared deviation from the mean**

The following is in **BOLD** to emphasize one of the most fundamental concepts in statistics, namely **VARIANCE**.

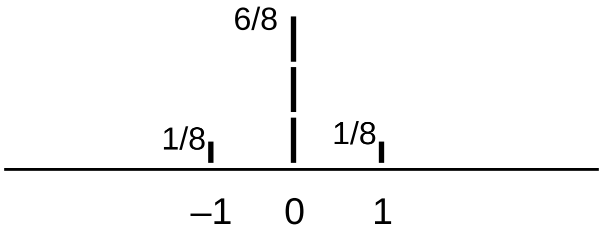
$E[(Y - \mu)^2]$ is called the variance of Y, usually shortened to Var(Y) or even to V(Y)

It, and its positive square root, called the standard deviation of Y, or SD(Y), are two of the most commonly used measures of variability or spread or uncertainty.

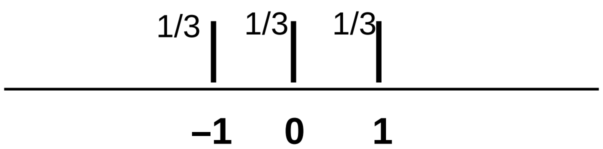
Although we first define variance and then take the square root to reach the SD, we should think of the SD as primary, at least for descriptive purposes (Mosteller et al use the natural order "...standard deviation and variance (either of these measures determines the other because the variance is the square of the SD"... cf. the average female fertility and the "variance" of this fertility.) However, there are good mathematical reasons to work with variance.

[Borrowing heavily from Mosteller et al. here...]
On the next page are 6 random variables, all symmetrical about the value y=0, so their means are all equal to zero. We consider various ways of measuring their spreads about this common mean. *(MRT2 p205)*

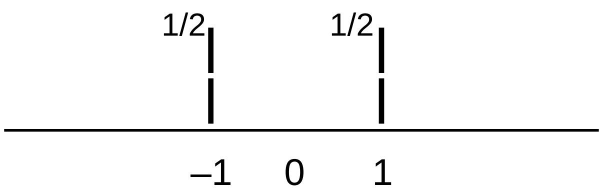
Y(A)



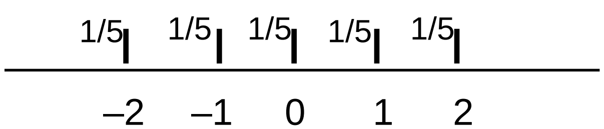
Y(B)



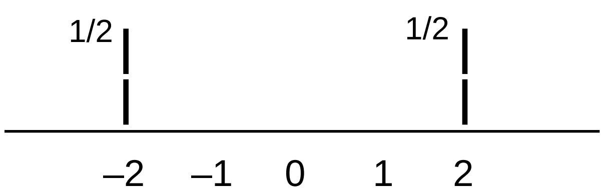
Y(C)



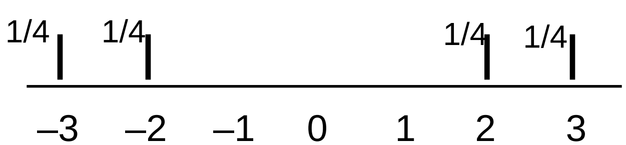
Y(D)



Y(E)



Y(F)



	Mean Absolute Deviation	**Variance**	**Standard Deviation**
Y(A)	$\frac{1}{4}$	$\frac{1}{4}$	**0.500**
Y(B)	$\frac{2}{3}$	$\frac{2}{3}$	**0.816**
Y(C)	**1**	**1**	**1.000**
Y(D)	$\frac{6}{5}$	**2**	**1.414**
Y(E)	**2**	**4**	**2**
Y(F)	$\frac{5}{2}$	$\frac{13}{2}$	**2.550**

TWO FUNDAMENTAL REASONS FOR USING VARIANCE, WHICH AVERAGES THE SQUARES OF DEVIATIONS FROM THE MEAN TO MEASURE VARIABILITY (From MRT2 p 207)

ADDITIVITY

The variance of the sum of two independent random variables is the sum of their variances, and even when the two variables are dependent the variability of their sum has a simple formula.

2. CENTRAL LIMIT THEOREM

The limiting behavior of a random variable that is the sum of a large number of independent random variables depends on the variances of these random variables.

DEFINITION: Variance

$$\text{Var}(Y) = E[(Y - \mu)^2] = \Sigma_y (y - \mu)^2 f(y)$$

DEFINITION: Standard Deviation

$$\sigma_Y = \sqrt{\text{Var}(Y)} = \sqrt{E[(Y - \mu)^2]}$$

USEFUL THEOREM (6.4, MRT2; 3.5 WMS5)

$$\text{Var}(Y) = E(Y^2) - \mu^2$$

Variance

=" average squared minus squared average"