Université Concordia University

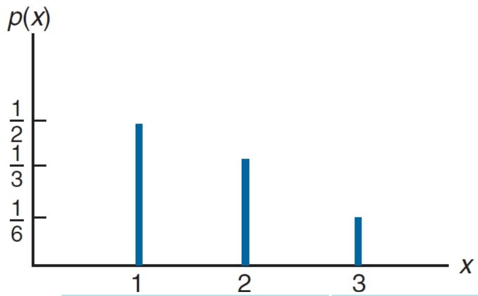
COMP 233/2
Probability and Statistics for Computer Science
Week 4

Expectation
Properties of Expectation
Variance
Covariance
Chebyshev's inequality

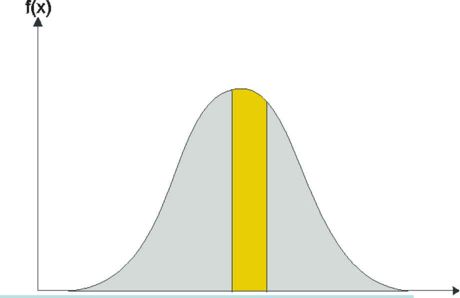
Reading: Chap 11

Information About Calculators During the Exams

- Yes, you may need a calculator.
- ONLY those calculators approved by the Exams Office — an ENCS sticker must be attached to the valid calculator, clearly indicating approval (during the final , exams invigilators will be checking).



Summary of Week 3



	Discrete RV	Continuous RV
PMF/ PDF	$p(x) = \begin{cases} p_i \geq 0, \text{ if } x = x_i \\ 0, \text{ otherwise} \end{cases}; \sum_{\text{all } x} p(x) = 1$	$f(x) \geq 0; \int_{-\infty}^{\infty} f(x)dx = 1.$
CDF	$F(a) = \sum_{\text{all } x \leq a} p(x)$	$F(a) = \int_{-\infty}^{a} f(x)dx.$
Limit	$F(\infty) = 1$	$F(\infty) = 1$
Intervals	$P\{a \leq X \leq b\} = \sum_{a \leq x \leq b} p(x)$ $= F(b) - F(a) + p(a)$	$P\{a \leq X \leq b\} = \int_{a}^{b} f(x)dx$ $= F(b) - F(a)$

Reminder: Independence

- In general, for two events, *E* and *F*, we have

$$P(EF) = P(E \mid F) \cdot P(F).$$

- If the events are, however, *independent*, then probability of the intersection is

$$P(EF) = P(E) \cdot P(F).$$

Independent Random Variables

- We call RV X and Y ***independent***, if for any two sets of real numbers A and B,

$$P\{X \in A, Y \in B\} = P\{X \in A\}P\{Y \in B\}.$$

- This implies that

$$F(x, y) = F_X(x) \cdot F_Y(y), \quad x \in R, y \in R.$$

$$p(x, y) = p_X(x) \cdot p_Y(y) \text{ (discrete case)},$$

$$\text{where } p_X \text{ and } p_Y \text{ are PMFs of } X \text{ and } Y.$$

$$f(x, y) = f_X(x) \cdot f_Y(y) \text{ (continuous case)}.$$

"Average" of a Discrete Random Variable

- Each time a random experiment is carried out, an associated RV takes on a value. Is it possible to find an "average" value of the RV ?
  - For instance, on average, how many bits are received in error when 8 bits are transmitted?
  - Or, on average, how many bits must be received until the first (the r-th) error occurs?
- Note that in both cases, you can take a sample, i.e. repeat the respective experiment several times and take the value of the RV each time.
- An average computed this way is, however, dependent on the sample!
- Is it possible to get a value independent of sampling?

Expectation

Expectation, or Expected Value

The ***expected value***, or ***expectation***, of a *discrete random variable* X is given by

$$\mu = E[X] = \sum_k x_k P\{X = x_k\} = \sum_k x_k p(x_k).$$

The formula gives a *weighted* average of values of a random variable, with the probabilities being the *weights* of the respective values. The value μ is termed the *mean of* X.

Example

- Find $E[X]$ where X is the outcome when we roll a fair die.

- Since $p(1) = p(2) = p(3) = p(4) = p(5) = p(6) = 1/6$, we obtain that

  $$E[X] = 1(1/6)+2(1/6)+3(1/6)+4(1/6)+5(1/6)+6(1/6)$$
  $$= 7/2 = 3.5$$

- Note that, for this example, the expected value of X is not a value that X could possibly assume.

Expectation for Continuous Density Function

- Similar to that for discrete RV, we can define the expectation for a continuous RV as

$$\mu = E[X] = \int_{-\infty}^{\infty} x f(x) dx.$$

- Thus, the $\sum$ is replaced by an $\int$

Example

- Suppose that you are expecting a message at some time past 5 P.M. From experience you know that X, the number of hours after 5 P.M. until the message arrives, is a random variable with the following probability density function:

$$f(x) = \begin{cases} \dfrac{1}{1.5} & \text{if } 0 < x < 1.5 \\ 0 & \text{otherwise} \end{cases}$$

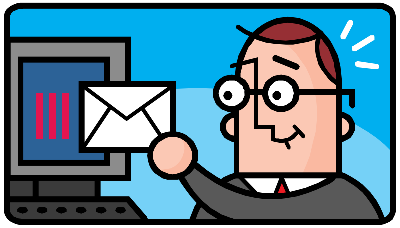
- What type of density function is this?

Solution

- The expected amount of time past 5 P.M. until the message arrives is given by

$$E[X] = \int_{-\infty}^{\infty} xf(x)dx = \left(\frac{1}{1.5}\right)\int_{0}^{1.5} xdx = \left(\frac{1}{1.5}\right)\left(\frac{x^2}{2}\Big|_0^{1.5}\right) = \frac{1.5}{2} = .75$$

- Hence, on average, you would have to wait three-fourths of an hour.



Expectation, or Expected Value

The ***expected value***, or ***expectation,*** of a *random variable* X is given by

$$\mu = E[X] = \begin{cases} \sum_k x_k p(x_k), & \text{discrete;} \\ \int x f(x)\,dx, & \text{continuous.} \end{cases}$$

Let's look at some (perhaps useful) properties of the expectation…

Scaling of Expectation

- Let a be a real number.
- Then aX is a new random variable with the same distribution as X.
- We observe that

$$E[aX] = \begin{cases} \int_{-\infty}^{\infty} axf(x)dx = a\int_{-\infty}^{\infty} xf(x)dx \\ \\ \sum_k ax_k p(x_k) = a\sum_k x_k p(x_k) \end{cases} = aE[X].$$

Scaling of Expectation, again

- Let a and b be real numbers.
- Then $aX + b$ is a new random variable with the same distribution as X.
- We observe that

$$E[aX+b] = \begin{cases} \displaystyle\int_{-\infty}^{\infty} (ax+b)f(x)dx = a\int_{-\infty}^{\infty} xf(x)dx + b\int_{-\infty}^{\infty} f(x)dx \\ \displaystyle\sum_k (ax_k+b)p(x_k) = a\sum_k x_k p(x_k) + b\sum_k p(x_k) \end{cases}$$

$$= aE[X] + b.$$

Example

- Suppose Y has the following probability mass function (PMF) $p(258) = .25$, $p(508) = .15$, $p(758) = .6$ Calculate $E[Y]$.
- **Solution** Letting $Y = 250X+8$, we have that X is a random variable that can take on one of the values 1, 2, 3 with respective probabilities

$$p_X(1) = P\{X = 1\} = .25$$
$$p_X(2) = P\{X = 2\} = .15$$
$$p_X(3) = P\{X = 3\} = .6$$

- Hence, $E[X] = 1(.25) + 2(.15) + 3(.6) = 2.35$
- Then $E[Y] = 250E[X]+8 = 595.5$

Université Concordia University

Example

- Suppose X has the following probability mass function $p(0) = .2$, $p(1) = .5$, $p(2) = .3$
  Calculate $E[X^2]$.
- **Solution** Letting $Y = X^2$, we have that Y is a random variable that can take on one of the values $0^2, 1^2, 2^2$ with respective probabilities

$$p_Y(0) = P\{Y = 0^2\} = .2$$
$$p_Y(1) = P\{Y = 1^2\} = .5$$
$$p_Y(4) = P\{Y = 2^2\} = .3$$

- Hence, $E[X^2] = E[Y] = 0(.2) + 1(.5) + 4(.3) = 1.7$

Joint Cumulative Distribution Function

- We introduce the ***joint cumulative distribution function*** for RV X and Y by

$$F(x, y) = P\{X \le x, Y \le y\}.$$

- F is non-negative.
- In addition, $F(x,y) \rightarrow F(x)$, as y approaches infinity, and $F(x,y) \rightarrow F(y)$, as x approaches infinity.

Example

- Suppose that 3 batteries are randomly chosen from a group of 3 new, 4 used but still working, and 5 defective batteries.
- If we let X denote the number of new batteries and Y the number of used but still working batteries that are chosen, then the joint probability mass function p of X and Y,

$$p(i, j) = P\{X = i, Y = j\},$$

is given by…

Example

$$p(0,0)=\binom{5}{3}\Big/\binom{12}{3}=\frac{10}{220}$$

$$p(0,1)=\binom{4}{1}\binom{5}{2}\Big/\binom{12}{3}=\frac{40}{220}$$

$$p(0,2)=\binom{4}{2}\binom{5}{1}\Big/\binom{12}{3}=\frac{30}{220}$$

$$p(0,3)=\binom{4}{3}\Big/\binom{12}{3}=\frac{4}{220}$$

$$p(1,0)=\binom{3}{1}\binom{5}{2}\Big/\binom{12}{3}=\frac{30}{220}$$

$$p(1,1)=\binom{3}{1}\binom{4}{1}\binom{5}{1}\Big/\binom{12}{3}=\frac{60}{220}$$

$$p(1,2)=\binom{3}{1}\binom{4}{2}\Big/\binom{12}{3}=\frac{18}{220}$$

$$p(2,0)=\binom{3}{2}\binom{5}{1}\Big/\binom{12}{3}=\frac{15}{220}$$

$$p(2,1)=\binom{3}{2}\binom{4}{1}\Big/\binom{12}{3}=\frac{12}{220}$$

$$p(3,0)=\binom{3}{3}\Big/\binom{12}{3}=\frac{1}{220}$$

Example

- Or in table form:

p(1,2)

used but still working

new

i \ j	0	1	2	3	P{X=i}
0	10/220	40/220	30/220	4/220	84/220
1	30/220	60/220	18/220	--	108/220
2	15/220	12/220	--	--	27/220
3	1/220	--	--	--	1/220
P{Y=j}	56/220	112/220	48/220	4/220	220/220

- The PMFs for X are obtained by the row sums, and the PMFs for Y are obtained by the column sums.

Example, cont.

- For the previous Joint Probability Mass Function for the batteries:
  a) Determine $F(2, 1)$.
  b) Determine $F_Y(1)$.
  c) What is the probability of 2 or more new batteries?

Joint Probability Density Function

- If the RV X and Y are **continuous**, then

$$F(s,t) = P\{X \le s, Y \le t\}$$

$$= \int_{-\infty}^{t}\int_{-\infty}^{s} f(x,y)\,dx\,dy.$$

- The function f is the ***joint probability density function*** of X and Y.
- Needless to say, the integration of f over the whole x-y plane equals 1 ($F(\infty,\infty) = 1$).

Development of Density for Sum.

- We shall compute the CDF of X from F(x,y).

$$
\begin{aligned}
F_X(t) &= P\{X < t\} \\
&= P\{X < t, -\infty < Y < \infty\} \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{t} f(x, y)dxdy \\
&= \int_{-\infty}^{t}\left(\int_{-\infty}^{\infty} f(x,y)dy\right)dx \\
&= \int_{-\infty}^{t} f_X(x)dx
\end{aligned}
$$

Example 1

- Let X stand for the time until a server connects to your machine, and let Y denote the time until the server authorizes you as a valid user. Each of the RV measures the wait from a common starting point. It is clear that $X < Y$.
- Under certain assumptions,

$$f(x, y) = \begin{cases} 0.000006\, e^{-0.001x - 0.002y} & 0 < x < y \\ 0 & \text{otherwise} \end{cases}$$

Example 1 (modified)

- Let X stand for the time until a server connects to your machine, and let Y denote the time from the moment of connection until the authorization as a valid user is complete. It is clear that now X > 0 and Y > 0.
- Under certain assumptions,

$$f(x,y) = \begin{cases} 0.000002\, e^{-0.001x-0.002y} & 0 < x, 0 < y \\ 0 & \text{otherwise} \end{cases}$$

Expectation for Joint Probability Distributions

- Consider a joint probability distribution of random variables X and Y. Then

$$E[X \cdot Y] = \begin{cases} \sum_{\text{all } x,y} xyp(x,y), & \text{discrete;} \\ \iint xyf(x,y)\,dx\,dy, & \text{continuous.} \end{cases}$$

- If X=Y, then this reduces to

$$E[X \cdot X] = \begin{cases} \sum_{\text{all } x,x} xxp(x,x) = \sum_{\text{all } x} x^2 p(x), & \text{discrete;} \\ \iint xxf(x,x)\,dx\,dx = \int x^2 f(x)\,dx, & \text{continuous.} \end{cases}$$

Example

- Consider the following joint PMF:

p(1,2)

i \ j	j=0	1	2	3	P{X=i}
i=0	0.1	0.1	0.05	0.05	0.3
1	0.1	0.1	0.05	0.05	0.3
2	0.1	0.05	0.025	0.025	0.2
3	0.1	0.05	0.025	0.025	0.2
P{Y=j}	0.4	0.3	0.15	0.15	1

- Determine $E[X], E[Y], E[X^2], E[Y^2]$, and $E[XY]$.

Problem

- Consider a pay-to-play game which involves flipping a coin three (3) times. The payout for the game depends on the number of heads obtained in the three coin flips. Let the discrete random variable X represent the number of heads.
  a) What is the probability $P\{X = k\}$ associated with each value k of the random variable?

Problem, cont.

- Suppose that the game has a payout of X^2 dollars. What is the minimum amount that should be charged for admittance (player gets one game of three coin flips) so that the person running the game won't lose money on average? Hint, determine $E[X^2]$.

Problem, cont.

- Suppose that the game has a new payout of $2X + 3X^2$ dollars. Now what is the minimum amount that should be charged for admittance?

Fundamental Properties

- Let X and Y be ***independent*** RVs. Then

$$E[X+Y] = E[X] + E[Y],$$
$$E[X \cdot Y] = E[X] \cdot E[Y].$$

- For example, if for X and Y, independent RVs, E[X] = 4 and E[Y] = 5, then

$$E[X+Y] = E[X] + E[Y] = 4 + 5 = 9$$
$$E[X \cdot Y] = E[X]E[Y] = (4)(5) = 20$$

Fundamental Properties, cont.

- The first equality (for E[X + Y]) is valid for all RV.

$$
\begin{aligned}
E(X+Y) &= \sum_j \sum_k (x_j + y_k) P(X = x_j,\ Y = y_k) \\
&= \sum_j \sum_k x_j P(X = x_j,\ Y = y_k) + \sum_j \sum_k y_k P(X = x_j,\ Y = y_k) \\
&= \sum_j x_j P(X = x_j) + \sum_k y_k P(Y = y_k)\ .
\end{aligned}
$$

Fundamental Properties, cont.

- In general, for any n,

$$E[X_1 + X_2 + \cdots + X_n] = E[X_1] + E[X_2] + \cdots + E[X_n]$$

- If $E[X_i] = \mu$, then

$$E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i] = n\mu$$

Fundamental Properties, cont.

- The proof for the second property (for $E[X \cdot Y]$) is simple. Assume that both X and Y are continuous.
- Then

$$E[X \cdot Y] = \iint_{R\times R} xyh(x,y)dxdy$$

$$= \iint_{R\times R} xyf(x)g(y)dxdy = \int_R xf(x)dx \int_R yg(y)dy$$

$$= E[X] \cdot E[Y].$$

Moments of X

- The expected value of a random variable X , E[X], is also referred to as the mean or the first moment of X . The quantity $E[X^n]$, $n \geq 1$, is called the *n-th moment* of X. Note that

$$E[X^n] = \begin{cases} \sum_k x_k^n p(x_k), & \text{discrete;} \\ \int x^n f(x)dx, & \text{continuous.} \end{cases}$$

Example

- The time, in hours, it takes to locate and repair an electrical breakdown in a certain factory is a random variable— call it X — whose density function is given by

$$f_X(\mathrm{x}) = \begin{cases} 1, & 0 < \mathrm{x} < 1, \\ 0, & \text{otherwise} \end{cases}$$

- If the cost involved in a breakdown of duration x is x^3, what is the expected cost of such a breakdown?

Solution

- Let X^3 denote the cost, then the expected cost is

$$E[X^3] = \int_{-\infty}^{\infty} x^3 f(x)dx$$
$$= \int_0^1 x^3 dx$$
$$= \frac{x^4}{4}\Big|_0^1 = \frac{1}{4}$$

Variance

Variation of X

- Given a random variable X along with its probability distribution function, it would be extremely useful if we were able to summarize the essential properties of the mass/density function by certain suitably defined measures.
- One such measure would be E[X], the expected value of X . However, while E[X] yields the weighted average of the possible values of X, it does not tell us anything about the variation, or spread, of these values.

Variance

- The average spread of a discrete RV around its *mean* is the expectation of the square of the difference between X and its mean μ called the ***variance***:

$$\begin{aligned} Var(X) &= E[(X-\mu)^2] \\ &= E[X^2] - 2E[X]\mu + \mu^2 \\ &= E[X^2] - \mu^2 \\ &= E[X^2] - E[X]^2 \end{aligned}$$

Example

- Compute Var(X) when X represents the outcome when we roll a fair die.
- **Solution** Since P{X = i} = 1/6, i = 1,2,...,6, we obtain

$$E[X^2] = \sum_{i=1}^{6} i^2 P\{X = i\}$$

$$= 1^2\left(\frac{1}{6}\right) + 2^2\left(\frac{1}{6}\right) + 3^2\left(\frac{1}{6}\right) + 4^2\left(\frac{1}{6}\right) + 5^2\left(\frac{1}{6}\right) + 6^2\left(\frac{1}{6}\right)$$

$$= \frac{91}{6}$$

Example, cont.

- Hence, since it was previously shown that

$$E[X] = 7/2$$

we obtain

$$\begin{aligned} Var(X) &= E[X^2] - E[X]^2 \\ &= \frac{91}{6} - \left(\frac{7}{2}\right)^2 \\ &= \frac{35}{12} \end{aligned}$$

Example (exponential RV)

- Recall the ***PDF*** of the *exponential* RV with λ>0,

$$f(t) = \begin{cases} 0, & t \le 0, \\ \lambda e^{-\lambda t}, & t > 0. \end{cases}$$

- A direct computation yields

$$E[X] = \int_{-\infty}^{\infty} tf(t)dt = \int_{0}^{\infty} \lambda t e^{-\lambda t} dt$$

$$= \left(-te^{-\lambda t} - \frac{e^{-\lambda t}}{\lambda} \right)\Big|_0^{\infty} = \frac{1}{\lambda}.$$

Example (exponential RV)

- Similarly,

$$E[X^2] = \int_{-\infty}^{\infty} t^2 f(t)dt = \int_{0}^{\infty} \lambda t^2 e^{-\lambda t} dt = \frac{2}{\lambda^2}.$$

- So that

$$Var(X) = E[X^2] - (E[X])^2$$

$$= \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}.$$

- Hence, $E[X] = 1/\lambda$, $Var(X) = 1/\lambda^2$.

Expectation and Variance of Other Density Functions

$$E[X] = \frac{\lambda\sqrt{\pi}}{2}, Var(X) = \lambda^2(1 - \frac{\pi}{4}) \quad \text{Weibull}(k = 2),$$

$$E[X] = \frac{\alpha + \beta}{2}, Var(X) = \frac{(\beta - \alpha)^2}{12} \quad \text{Uniform.}$$

Scaling of Variance

- Let a be a real number. Then aX is a new random variable with the same distribution as X, and

$$Var(aX) = \int_{-\infty}^{\infty} (a\mathrm{x} - a\mu)^2 f(\mathrm{x}) d\mathrm{x}$$

$$= a^2 \int_{-\infty}^{\infty} (\mathrm{x} - \mu)^2 f(\mathrm{x}) d\mathrm{x} = a^2 Var(X).$$

- In particular, $E[X+X] = E[2X] = 2E[X]$, and

$$Var(X+X) = Var(2X) = 4Var(X) \neq 2\mathrm{Var}(\mathrm{X}).$$

Reminder: Intersection

- In general, for two events, *E* and *F*, we have

$$P(EF) = P(E \mid F) \cdot P(F).$$

- If the events are, however, *independent*, then probability of the intersection is

$$P(EF) = P(E) \cdot P(F).$$

Independent Random Variables

- We call RV X and Y ***independent***, if for any two sets of real numbers A and B,

$$P\{X \in A, Y \in B\} = P\{X \in A\}P\{Y \in B\}.$$

- This implies that

$$F(x, y) = F_X(x) \cdot F_Y(y), \quad x \in R, y \in R.$$

$$p(x, y) = p_X(x) \cdot p_Y(y) \text{ (discrete case)},$$

$$\text{where } p_X \text{ and } p_Y \text{ are PMFs of } X \text{ and } Y.$$

$$f(x, y) = f_X(x) \cdot f_Y(y) \text{ (continuous case)}.$$

Covariance

- The covariance of two random variables X and Y, written Cov(X, Y), is defined by

$$Cov(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$$

where μ_X and μ_Y are the means of X and Y, respectively.

- A useful expression for Cov(X, Y) is

$$Cov(X, Y) = E[XY] - E[X]E[Y]$$

- Note that

$$Cov(X,Y) = Cov(Y,X)$$

$$Cov(X,X) = Var(X)$$

Variance of Sum of RVs

- Let X and Y be RVs. Then

$$Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y).$$

- The covariance term provides a measure of the strength of the correlation between X and Y.
- In general, for n RVs,

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) + \sum_{i=1}^{n} \sum_{\substack{i=1 \\ i \neq j}}^{n} Cov(X_i, X_j).$$

Sum of Independent RVs

- Let X and Y be *independent* RVs. Then

$$Cov(X,Y)=0$$

which implies that

$$Var(X+Y)=Var(X)+Var(Y).$$

- In general, for n independent RVs, X_k, k$=1,\dots,n$,

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

Food for Thought - 4.2

- Consider the following joint PMF:

i \ j	j=0	1	2	3	P{X=i}
i=0	0.1	0.1	0.05	0.05	0.3
1	0.1	0.1	0.05	0.05	0.3
2	0.1	0.05	0.025	0.025	0.2
3	0.1	0.05	0.025	0.025	0.2
P{Y=j}	0.4	0.3	0.15	0.15	1

- **Determine** $Var(X), Var(Y), Cov(X,Y)$, **and** $Var(X+Y)$.

Standard Deviation

- The quantity $$\sqrt{Var(X)}$$ is called the standard deviation of X .
- The standard deviation has the same units as does the mean μ.

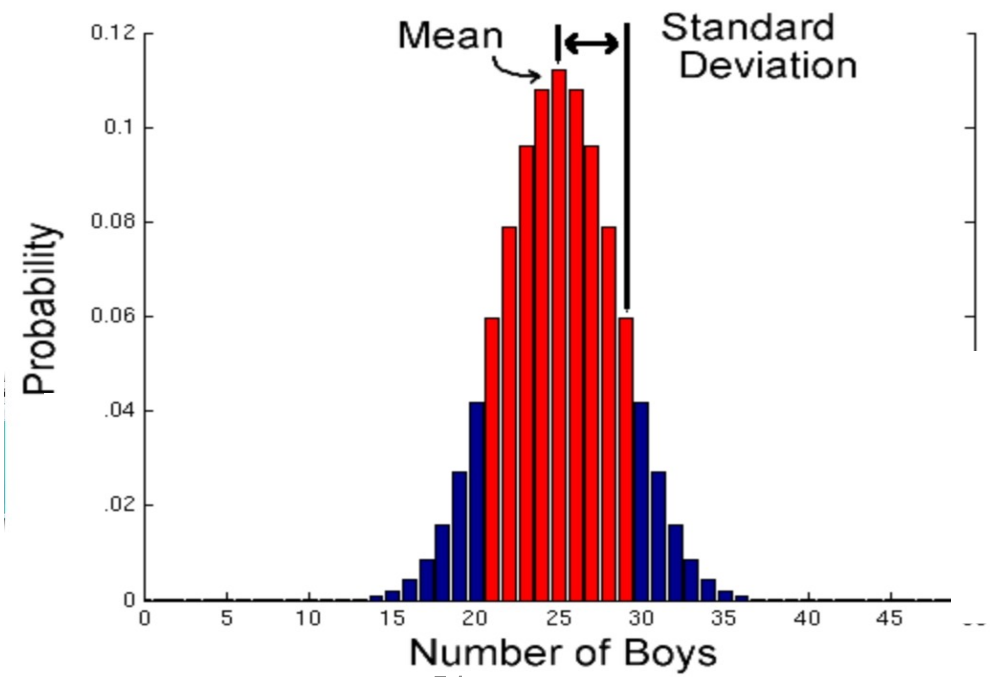


Figure source: http://en.wikipedia.org/wiki/File:The-giver-binomial.png

Correlation

- A positive value of Cov(X , Y) is an indication that Y tends to increase as X does, whereas a negative value indicates that Y tends to decrease as X increases.
- The strength of the relationship between X and Y is indicated by the correlation between X and Y:

$$Corr(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}.$$

- The correlation has a value between -1 and 1.

Summary

- Let X_k, k$=1,\ldots,n,$ be *independent* RV. Then

$$E[X_1 + \ldots + X_n] = \sum_k E[X_k],$$

$$Var(X_1 + \ldots + X_n) = \sum_k Var(X_k).$$

  - It remains to discuss the *distribution* of the sum (i.e., is it exponential, uniform, something else?), which will be a later topic of discussion.

Notes on Discrete RV

- We notice that
  - The discrete Probability Mass Functions can be described by their CDFs;
  - The CDFs are non-decreasing step functions;
  - The CDFs only make jumps at points in the range of the random variable.
  - One can recover the respective mass functions using the size of the jump;
  - The expectation and the variance are computed using the respective mass function;
  - The calculation of probabilities, expectation and variance involves summation.

Notes on Continuous RV

- We notice that
  - The continuous Probability Density Functions are described by their CDF in a way similar to discrete RV;
  - The CDF are continuous;
  - The derivatives of CDF – the density functions – are similar to mass functions for discrete mass functions;
  - The expectation and variance are computed via certain integrals of PDF;
  - Whenever we used summation in discrete case, we had to integrate in the continuous case.

Chebyshev's inequality

Markov's Inequality

- If X is a random variable that takes only nonnegative values, then for any value $a > 0$, an upper bound on the probability is

$$P\{X \geq a\} \leq \frac{E[X]}{a}$$

- The corresponding lower bound is

$$P\{X < a\} \geq 1 - \frac{E[X]}{a}$$

Chebyshev's Inequality

- If X is a random variable with mean μ and variance σ^2 = Var(X), then for any value k > 0, an (usually) improved upper bound on the probability is

$$P\{|X-\mu| \geq k\} \leq \frac{\sigma^2}{k^2}$$

Deriving Bounds on Probabilities

- The importance of Markov's and Cheybyshev's inequalities is that they enable us to *derive bounds* on probabilities when only the mean, or both the mean and the variance, of the probability distribution are known.
- Of course, if the actual distribution were known, then the desired probabilities could be exactly computed and we would not need to resort to these rough bounds.

Food for Thought – 4.2

- Suppose that it is known that the number of items produced in a factory during a week is a random variable with mean 50.

  (a) What can be said about the probability that this week's production will *exceed* 75?

  (b) If the variance of a week's production is known to equal 25, then what can be said about the probability that this week's production will be *between* 40 and 60?

Université Concordia University

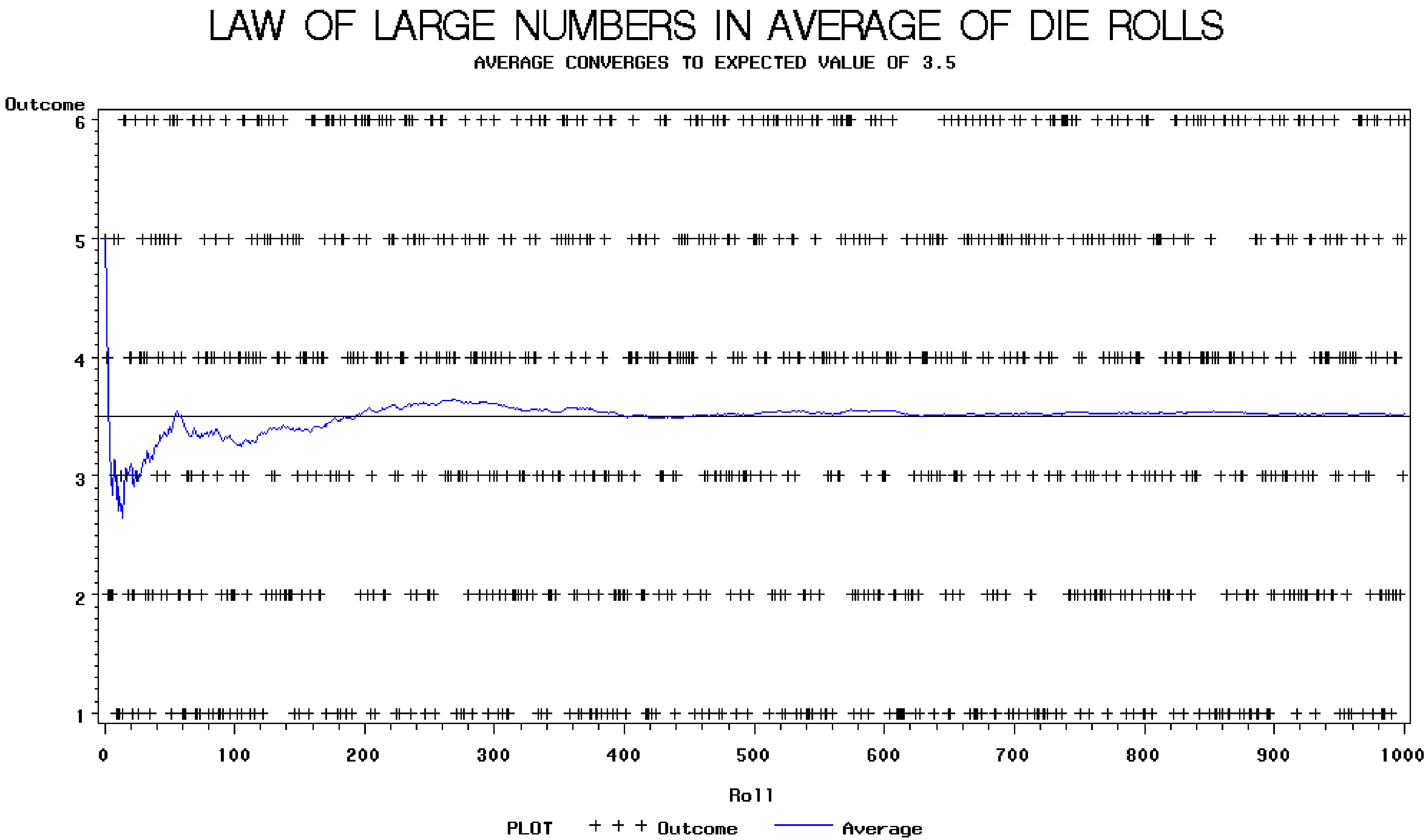
The Weak Law of Large Numbers

- Let $X_1, X_2, \ldots,$ be a sequence of independent and identically distributed random variables, each having mean $E[X_i] = \mu$. Then, for any $\varepsilon > 0$,

$$P\left\{\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| > \varepsilon\right\} \to 0 \text{ as } n \to \infty$$

- In other words, for any positive number ε, no matter how small, the probability that the proportion of the first n trials in which an event occurs differs from the mean by more than ε goes to 0 as n increases.

Illustration



Source: http://en.wikipedia.org/wiki/Law_of_large_numbers

Food for Thought - 4.3

- We can use the Chebyshev inequality to prove the Weak Law of Large Numbers.
- That is use the Chebyshev inequality,

$$P\{|X-\mu| \geq k\} \leq \frac{Var(X)}{k^2}$$

And prove:

$$P\left\{\left|\frac{X_1+\cdots+X_n}{n}-\mu\right| > \varepsilon\right\} \to 0 \text{ as } n \to \infty$$

Special Random Variables – next class

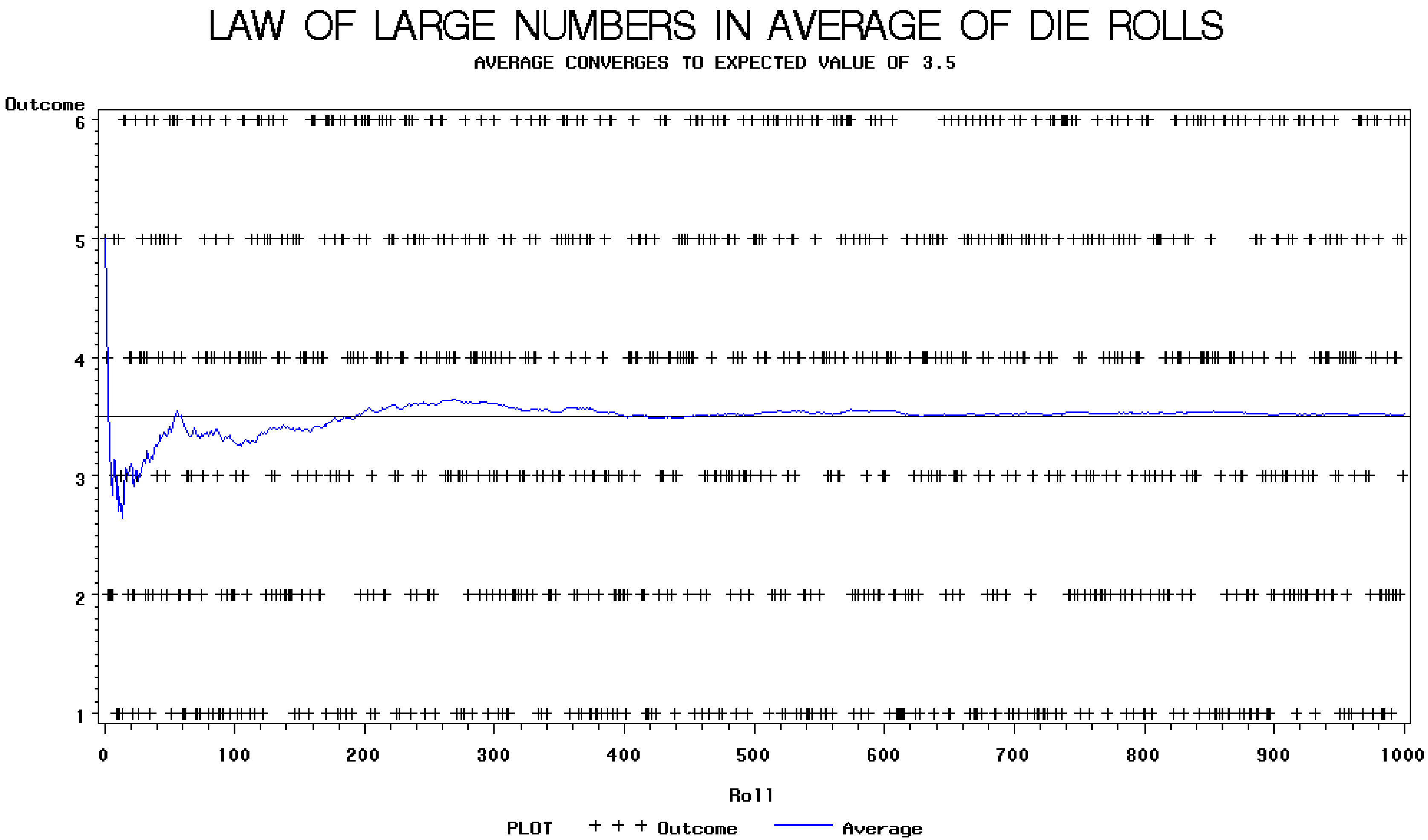
- The properties like expectation and variance allow us to compute general bounds on probabilities, but to get more exact answers, we need to know the probability distribution of the random variable.
- Next, we will study certain types of random variables that occur over and over again in applications.
- So next class' topic is

<u>SPECIAL RANDOM VARIABLES</u>

References/Resources Used

- Lecture Slides for MATH 401 of Dr. Oleksiy Us, Department of Mathematics, German University of Cairo. [PPT]

Illustration



Source: http://en.wikipedia.org/wiki/Law_of_large_numbers