Special Mathematics
Discrete random variables

April 2018
“Expose yourself to as much randomness as possible.”

Ben Casnocha

6

Discrete random variables

Texas Holdem Poker:

In Hold’em Poker players make the best hand they can combining the two cards in their hand with the 5 cards (community cards) eventually turned up on the table. The deck has 52 and there are 13 of each suit: ♠ ♡ ♥ ♣.

Open problem: Until he finally won a WSOP event in 2008, Erick Lindgren was often called one of the greatest players never to have won a WSOP tournament. Before his win, he played in many WSOP events and finished in
the top 10 eight times. Suppose you play in one tournament per week. For simplicity, assume that each tournament’s results are independent of the others and that you have the same probability \( p \) of winning each tournament. If \( p = 0.01 \), then what is the expected amount of time before you win your first tournament?

Open problem: During Episode 2 of Season 5 of High Stakes Poker, Doyle Brunson was dealt pocket kings twice and pocket jacks once, all within about half an hour. Suppose we consider a high pocket pair to mean 10-10, J-J, Q-Q, K-K, or A-A. Let \( X \) be the number of hands you play until you are dealt a high pocket pair for the third time. What is the expected number of hands?

Open problem: Many casinos award prizes for rare events called jackpot hands. These jackpot hands are defined differently by different casinos. Suppose in a certain casino jackpot hands are defined so that they tend to occur about once every 50,000 hands on average. If the casino deals about 10,000 hands per day, what are the expected value and standard deviation of the number of jackpot hands dealt in a 7-day period?

Open problem: On the last hand of the 1998 WSOP Main Event, with the board \( 8 \spadesuit 9 \spadesuit 9 \heartsuit 8 \spadesuit 8 \spadesuit \), Scotty Nguyen went all-in. While his opponent, Kevin McBride, was thinking, Scotty said, “You call, it’s gonna be all over, baby.” McBride said, “I call. I play the board.” It turned out that Scotty had \( J \diamond 9 \spadesuit \) and won the hand. Assuming you never fold in the next 100 hands, what would be the expected value of \( X = \text{the number of times in these 100 hands that you would play the board after all five board cards are dealt} \)?
Discrete random variables

Bernoulli's random variable \( X \sim \text{Ber}(p) \)

- it is the simplest discrete random variable, it models an experiment in which only two possible outcomes can occur, often designated success, and failure.

\[ X : \begin{pmatrix}
0 & 1 \\
1 - p & p \\
\end{pmatrix} \]

Example:

It can be used to represent a coin toss (we consider the appearance of a tail as being a success). We assign the value 1 to success with the probability \( p \in (0, 1) \) and the value 0 to failure with probability \( q = 1 - p \). Thus we obtain a Bernoulli random variable \( X \) with parameter \( p \) (probability of a success) \( X \sim \text{Ber}(p) \):

\[ X : \begin{pmatrix}
0 & 1 \\
1 - p & p \\
\end{pmatrix} \]

Of course, in our example \( p = \frac{1}{2} \). □

Uniform discrete random variable \( X \sim \mathcal{U}(n) \)

It represents a mathematical model which generalizes the experiment of throwing a die (case \( n = 6 \))

- if an experiment has \( n \) equally possible outcomes denoted \( \{1, 2, \ldots, n\} \), then the experiment can be modelled using a uniform random variable of the form:

\[ X : \begin{pmatrix}
1 & 2 & \ldots & n \\
\frac{1}{n} & \frac{1}{n} & \ldots & \frac{1}{n} \\
\end{pmatrix} \]

Binomial random variable \( X \sim \text{Bin}(n,p) \)

A random variable with a binomial distribution is the right model when the following assumptions hold (we have a binomial experiment):

- the modelled phenomenon consists of \( n \) independent trials of the same experiment
- there are only two possible outcomes at each trial (success - failure)
- the probability \( p \) of a success is the same at each trial

The random variable which counts the number of successes in \( n \) trials of a binomial experiment is called a binomial random variable:

\[ X : \begin{pmatrix}
0 & 1 & \ldots & k & \ldots & n \\
q^n & C_n^1pq^{n-1} & \ldots & C_n^kpq^{n-k} & \ldots & p^n \\
\end{pmatrix} \]
where $p$ and $q = 1 - p$ are the probabilities of a success, respectively failure at each independent trial.

**Geometric random variable $X \sim Geo(p)$**

It is the appropriate model if in a binomial experiment we count the number of failures until the first success occurs.

$$X : \begin{pmatrix} 0 & 1 & \ldots & k & \ldots \\ p & p(1-p) & \ldots & p(1-p)^k & \ldots \end{pmatrix}$$

**Hypergeometric random variable $X \sim Hip(N, M, n)$**

Consider the problem of drawing objects from a box which contains $N$ objects, with $M$ of them being defective.

- if the draws are with replacement (the extracted object is put back in the box before the next draw), then the number of the defective objects drawn in $n$ draws is a binomial random variable with parameters $n$ and $p = \frac{M}{N}$ (probability of drawing a defective object)

- if the draws are without replacement, then the probability to draw a defective object is not the same in each of those $n$ draws, thus the number of the defective objects drawn is no more a binomial random variable.

$\implies$ one obtains a random variable having the probability mass function

$$P(X = k) = \begin{cases} \frac{\binom{k}{M} \binom{N-k}{n-k}}{\binom{N}{n}}, & \text{if } k \in \{0, 1, 2, \ldots, n\} \\ 0, & \text{otherwise} \end{cases}$$

and it is called a hypergeometric random variable with parameters $M, N$ and $n$.

**Poisson’s random variable $X \sim Po(\lambda)$**

First of all, according to Poisson’s law:

$$P(X = k) = \binom{n}{k} p^k q^{n-k} \approx \frac{\lambda^k}{k!} e^{-\lambda}, \quad \text{for } \lambda = np$$

we can approximate the distribution series of a binomial distribution when the probability $p$ of a success at each trial is small and the number of trials $n$ is big. In practice we usually apply it for $p < 0.05$ and $n \geq 100$

- this law generates a probability distribution that expresses the probability of a given number of events occurring in a fixed interval of time or space if these events occur with a known constant rate and independently of the time since the last event. The Poisson distribution can also be used for the number of events in other specified intervals such as distance, area or volume.

- the Poisson distribution is usually used for rare events and it is also called the law of rare events.

$$X : \begin{pmatrix} 0 & 1 & \ldots & k & \ldots \\ e^{-\lambda} & \frac{\lambda}{1!} e^{-\lambda} & \ldots & \frac{\lambda^k}{k!} e^{-\lambda} & \ldots \end{pmatrix}$$
Negative binomial random variable $X \sim NB(r, p)$

- is a discrete probability distribution of the number of successes in binomial experiment before a specified number of failures denoted $r$ occurs.

$$X : \begin{pmatrix} 0 & 1 & \ldots & k & \ldots \\ (1 - p)^r & C_{1+r-1}^1 p(1 - p)^r & \ldots & C_{k+r-1}^k p^k(1 - p)^r & \ldots \end{pmatrix}$$

- sometimes we want to count the number of trials needed to produce the $r$-th success

such a random variable $X$ will have the probability mass function:

$$P(X = k) = C_{r-1}^{r-1} p^r (1 - p)^k, \quad k = r, r + 1, r = 2, \ldots$$

The above identity is interpreted as: "the probability to obtain in the $k$-th trail the $r$-th success is..."

**Numerical characteristics of random variables**

- for a discrete random variable:

$$X : \begin{pmatrix} x_i \\ p_i \end{pmatrix} \quad i \in I$$

the expected valued or mean $E(X)$ or $M(X)$ is defined as:

$$\bar{x} = E(X) = M(X) = \sum_{i \in I} p_i x_i$$

- the moments $M_k$ and the central moments $m_k$ of order $k$ are defined by:

$$M_k(X) = M(X^k) = \sum_{i \in I} p_i x_i^k,$$

$$m_k(X) = M((X - \bar{x})^k) = \sum_{i \in I} p_i (x_i - \bar{x})^k$$

- the variance is defined as:

$$\text{var}(X) = \sum_{i \in I} p_i (x_i - \bar{x})^2$$

and the standard deviation is:

$$\sigma(X) = \sqrt{\text{var}(X)}$$

- for $X \sim Bin(n, p) \implies E(X) = np$ and $\text{var}(X) = np(1 - p)$
- for $X \sim Ber(p) \implies E(X) = p$ and $\text{var}(X) = p(1 - p)$
- for $X \sim Geo(p) \implies E(X) = \frac{1}{p}$ and $\text{var}(X) = \frac{1-p}{p^2}$
- for $X \sim Po(\lambda) \implies E(X) = \lambda$ and $\text{var}(X) = \lambda$
- for $X \sim NB(r, p) \implies E(X) = \frac{r}{p}$ and $\text{var}(X) = \frac{r(1-p)}{p^2}$
• the following identities hold:

\[ E(X + aY) = E(X) + aE(Y) \]
\[ \text{var}(aX + C) = a^2 \text{var}(X) \]

where \( C \) is the constant random variable.

• if \( X \) and \( Y \) are independent:

\[ E(XY) = E(X)E(Y) \]
\[ \text{var}(X + aY) = \text{var}(X) + a^2 \text{var}(Y). \]
Solved problems

**Problem 1.** Three shooters shoot a target. The random variable $X$ which counts the number of hits has the distribution series

$$X = \begin{pmatrix} 0 & 1 & 2 & 3 \\ \frac{p^2}{4} & \frac{11p}{24} & \frac{1}{4} & \frac{1}{24} \end{pmatrix}.$$ 

a) After one finds the value of $p$, compute the probability that $X$ takes a value smaller or equal with 2.
b) Find the probability of hitting the target for each shooter.

**Solution:**

a) The sum of all probabilities in a distribution series of a random variable must be 1, thus

$$\frac{p^2}{4} + \frac{11p}{24} + \frac{1}{4} + \frac{1}{24} = 1 \iff 6p^2 + 11p - 17 = 0 \Rightarrow p = 1$$

$$P(X \leq 2) = 1 - P(X = 3) = 1 - P(X > 2) = 1 - \frac{1}{24} = \frac{23}{24}$$

b) Let $p_1, p_2, p_3$ be these probabilities. Hence we have for $p = 1$:

$$X = \begin{pmatrix} 0 & 1 & 2 & 3 \\ \frac{1}{4} & \frac{11}{24} & \frac{1}{4} & \frac{1}{24} \end{pmatrix}$$

But:

$$\frac{1}{4} = P(X = 0) = (1 - p_1)(1 - p_2)(1 - p_3)$$

(because $X = 0$ means: all the shooters missed the target)

$$= 1 - (p_1 + p_2 + p_3) + p_1p_2 + p_1p_3 + p_2p_3 - p_1p_2p_3$$

$$\frac{11}{24} = P(X = 1) = p_1(1 - p_2)(1 - p_3) + p_2(1 - p_1)(1 - p_3) + p_3(1 - p_1)(1 - p_2)$$

(because $X = 1$ means: one of the shooters hit and the others missed the target)

$$= p_1 + p_2 + p_3 - 2(p_1p_2 + p_1p_3 + p_2p_3) + 3p_1p_2p_3$$

$$\frac{1}{4} = P(X = 2) = p_1p_2(1 - p_3) + p_1p_3(1 - p_2) + p_2p_3(1 - p_1)$$

$$= p_1p_2 + p_1p_3 + p_2p_3 - 3p_1p_2p_3$$

$$\frac{1}{24} = P(X = 3) = p_1p_2p_3.$$
One gets the linear system
\[
\begin{align*}
\begin{cases}
p_1 + p_2 + p_3 &= \frac{13}{12} \\
p_1 p_2 + p_1 p_3 + p_2 p_3 &= \frac{3}{8} \\
p_1 p_2 p_3 &= \frac{1}{24}
\end{cases}
\end{align*}
\]
which leads to the equation
\[24x^3 - 26x^2 + 9x - 1 = 0\]
with the roots
\[p_1 = \frac{1}{2}; p_2 = \frac{1}{3}; p_3 = \frac{1}{4}.
\]

**Problem 2.** The random variables \(X\) and \(Y\) have the distribution series
\[
X = \begin{pmatrix} 1 & 2 & 3 \\ 0,1 & 0,2 & 0,7 \end{pmatrix}, \quad Y = \begin{pmatrix} 4 & 5 & 6 \\ 0,4 & 0,5 & 0,1 \end{pmatrix}.
\]
Compute:
\(\text{a) The distribution series of } X + Y,\)
\(\text{b) The distribution series of } X \cdot Y,\)
\(\text{c) The distribution series of } X^2.\)

**Solution:**
\(\text{a) The distribution series of } X + Y\) is:
\[
X + Y : \begin{pmatrix} 5 & 6 & 7 & 8 & 9 \\ 0,04 & 0,13 & 0,39 & 0,37 & 0,07 \end{pmatrix},
\]
For example when \(X + Y = 6\) then:
\[
P(X + Y = 6) = P(X = 1 \text{ if } Y = 5) + P(X = 2 \text{ and } Y = 4) \\
= P(X = 1) \cdot P(Y = 5) + P(X = 2) \cdot P(Y = 4) \\
(\text{the random variables are considered independent}) \\
= 0,1 \cdot 0,5 + 0,2 \cdot 0,4 = 0,05 + 0,08 = 0,13
\]
\(\text{b) The distribution series of } X \cdot Y\) is:
\[
X \cdot Y : \begin{pmatrix} 4 & 5 & 6 & 8 & 10 & 12 & 15 & 18 \\ 0,04 & 0,05 & 0,01 & 0,08 & 0,1 & 0,3 & 0,35 & 0,07 \end{pmatrix},
\]
For example when \(X \cdot Y = 4\):
\[
P(X \cdot Y = 4) = P(X = 1 \text{ and } Y = 4) \\
= P(X = 1) \cdot P(Y = 4) \\
= 0,1 \cdot 0,4 = 0,04
\]
again using the independence of \(X\) and \(Y\) we were able to compute straightforward 
\(P(X = 1 \text{ and } Y = 4) = P(X = 1) \cdot P(Y = 4)\).

c) For the random variable \(X^2\) the distribution series is:
\[
X^2 = \begin{pmatrix} 1 & 4 & 9 \\ 0 & 1 & 0 & 2 & 0 & 7 \end{pmatrix}.
\]

Generally for a function \(g\) the distribution series of \(Y := g(X)\) can be computed using the formula:
\[
P(Y = y) = \sum_{x: g(x) = y} P(X = x)
\]
adding all the probabilities \(P(X = x)\) for those \(x\) with the property \(g(x) = y\). In the above example the function \(g(x) = x^2\) is injective for positive values and the sum in that formula will contain only one term. Thus, for example:
\[
P(Y = 4) = \sum_{x: x^2 = 4} P(X = x) = P(X = 2) = 0, 2
\]

**Problem 3.** There are 3 traffic barriers along a street. The probability that a car which drives along that street finds any of these three barriers open is \(p = 0.8\). We suppose that any of these barriers work independently. Compute:
a) The distribution series of the random variable which counts the number of barriers passed until the first closed barrier met.
b) Find its cumulative distribution function.
c) Which is the expected number of barriers found open before the car has to stop in front of a closed one?

**Solution:**
a) We denote by \(X\) the desired random variable, which has a distribution series:
\[
X = \begin{pmatrix} 0 & 1 & 2 & 3 \\ p_0 & p_1 & p_2 & p_3 \end{pmatrix},
\]
where \(p_k = P(X = k), k = 0, 1, 2, 3\). By the way we defined the random variable one gets easily:
\[
p_0 = P(X = 0) = 0, 2
p_1 = P(X = 1) = 0.8 \cdot 0.2 = 0, 16
p_2 = P(X = 2) = 0.8 \cdot 0.8 \cdot 0.2 = 0, 128
p_3 = P(X = 3) = 0.8 \cdot 0.8 \cdot 0.8 = 0, 512
\]
Hence:
\[
X = \begin{pmatrix} 0 & 1 & 2 & 3 \\ 0, 2 & 0, 16 & 0, 128 & 0, 512 \end{pmatrix}.
\]
b) When \( x \leq 0 \) we get \( F(x) := P(X < x) = 0 \) because in the interval \((-\infty, 0)\) there are no values of \( X \).

When \( 0 < x \leq 1 \) one gets:

\[
F(x) = P(X < x) = P(X = 0) = 0, 2.
\]

When \( 1 < x \leq 2 \) one gets:

\[
F(x) = P(X < x) = P(X = 0) + P(X = 1)
\quad = 0, 2 + 0, 16 = 0, 36.
\]

When \( 2 < x \leq 3 \) one gets:

\[
F(x) = P(X < x) = P(X = 0) + P(X = 1) + P(X = 2)
\quad = 0, 2 + 0, 16 + 0, 128 = 0, 488.
\]

When \( x > 3 \) we have \( F(x) = 1 \).

Thus the cumulative distribution function of \( X \) is:

\[
F(x) = \begin{cases} 
0 & , \ x \leq 0 \\
0, 2 & , \ 0 < x \leq 1 \\
0, 36 & , \ 1 < x \leq 2 \\
0, 488 & , \ 2 < x \leq 3 \\
1 & , \ 3 < x 
\end{cases}
\]

Remark: Some authors define the cumulative distribution function as \( F(x) := P(X \leq x) \) then the above result looks differently but we think in a similar manner discussing the cases \( k \leq x < k + 1 \).

c) The driver expects to find 2 barriers open because the expected value of \( X \) is \( E(X) = 0 \cdot 0, 2 + 1 \cdot 0, 16 + 2 \cdot 0, 128 + 3 \cdot 0, 512 \approx 1.95 \).
Proposed problems

Problem 1. From a lot of 100 items, of which 10 are defective a random sample of size 5 is selected for quality control. Construct the distribution series of the random number $X$ of defective items contained in the sample.

Problem 2. A car has four traffic lights on its route. Each of them allows it to move ahead or stop with the probability 0.5. Sketch the distribution polygon of the probabilities of the numbers of lights passed by the car before the first stop has occurred.

Problem 3. Births in a hospital occur randomly at an average rate of 1.8 births per hour. What is the probability of observing 4 births in a given hour at the hospital?

Problem 4. It is known that 3% of the circuit boards from a production line are defective. If a random sample of 120 circuit boards is taken from this production line estimate the probability that the sample contains:

i) Exactly 2 defective boards.

ii) At least 2 defective boards.

Problem 5. Four different prizes are randomly put into boxes of cereal. One of the prizes is a free ticket to the local zoo. Suppose that a family of four decides to buy this cereal until obtaining four free tickets to the zoo. What is the probability that the family will have to buy 10 boxes of cereal to obtain the four free tickets to the zoo? What is the probability that the family will have to buy 16 boxes of cereal to obtain the four free tickets to the zoo?

Problem 6. An automatic line in a state of normal adjustment can produce a defective item with probability $p$. The readjustment of the line is made immediately after the first defective item has been produced. Find the average number of items produced between two readjustments of the line.

Problem 7. For two independent random variables:

\[
X:\begin{pmatrix} 0 & 1 & 2 & 3 \\ \frac{1}{8} & \frac{3}{8} & \frac{2}{8} & \frac{2}{8} \end{pmatrix}
\]

and:

\[
Y:\begin{pmatrix} 0 & 1 & 2 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}
\]

compute $X + Y$, $2X$, $M(X)$ and show that $M(XY) = M(X)M(Y)$. Compute $\text{var}(X + 2Y)$. 

11
Problem 8. A student takes a multiple-choice test consisting of two problems. The first one has 3 possible answers and the second one has 5. The student chooses, at random, one answer as the right one from each of the two problems. Find the expected number $E(X)$ of right answers $X$ of the student. Find the variance $\text{var}(X)$. Generalize.

Problem 9. The number of calls coming per minute into a hotel’s reservation center is a Poisson random variable of parameter $\lambda = 3$.

(a) Find the probability that no calls come in a given 1 minute period.

(b) Assume that the number of calls arriving in two different minutes are independent. Find the probability that at least two calls will arrive in a given two minute period.

(c) What is the expected number of calls in a given period of 1 minute?

Problem 10. As a result of experiments with two devices $A$ and $B$, one finds the probability of observing a noise whose level is evaluated in a three-point system:

<table>
<thead>
<tr>
<th>Noise level</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Device $A$</td>
<td>0.20</td>
<td>0.06</td>
<td>0.04</td>
</tr>
<tr>
<td>Device $B$</td>
<td>0.06</td>
<td>0.04</td>
<td>0.10</td>
</tr>
</tbody>
</table>

Using the table select the device with lower noise level.