N.I. Lobachevsky State University
of Nizhni Novgorod

Probability theory and mathematical statistics:

Dependence and independence of random variables Conditional distribution

Associate Professor
A.V. Zorine

Consider two random variables X and Y. Let numbers $x_1, x_2, \dots,$ be the possible values of X, and $y_1, y_2, \dots$ be the possible values of Y. Probabilities

$$p_{i,j} = \mathrm{P}(X = x_i, Y = y_j) = \mathrm{P}(\{\omega\colon X(\omega) = x_i\} \cap \{\omega\colon Y(\omega) = y_i\})$$

determine *the joint distribution* of random variables X and Y.
Any numbers $p_{i,j} \geqslant 0$ can assign probabilities to elements from the set $\{(x_i, y_j); i = 1, 2, \dots,\ j = 1, 2, \dots\}$ as soon as

$$\sum_{i \geqslant 1} \sum_{j \geqslant 1} p_{i,j} = 1$$

Example. Let a coin be tossed once. Let X be the number of Heads, Y the number of Tails. Then $\Omega = \{H, T\}$.

Example. Let a coin be tossed once. Let X be the number of Heads, Y the number of Tails. Then $\Omega = \{H, T\}$.

ω	$X(\omega)$	$Y(\omega)$
H	1	0
T	0	1

Example. Let a coin be tossed once. Let X be the number of Heads, Y the number of Tails. Then $\Omega = \{H, T\}$.

Now,

ω	$X(\omega)$	$Y(\omega)$
H	1	0
T	0	1

$$\{\omega \colon X(\omega) = 0, Y(\omega) = 0\} = \varnothing,$$

$$\mathrm{P}(X = 0, Y = 0) = 0$$

Example. Let a coin be tossed once. Let X be the number of Heads, Y the number of Tails. Then $\Omega = \{H, T\}$.

Now,

ω	$X(\omega)$	$Y(\omega)$
H	1	0
T	0	1

$$\{\omega\colon X(\omega) = 0, Y(\omega) = 0\} = \varnothing,$$

$$\mathrm{P}(X = 0, Y = 0) = 0$$

$$\{\omega\colon X(\omega) = 1, Y(\omega) = 0\} = \{H\},$$

$$\mathrm{P}(X = 1, Y = 0) = \frac{1}{2}$$

Example. Let a coin be tossed once. Let X be the number of Heads, Y the number of Tails. Then $\Omega = \{H, T\}$.

Now,

ω	$X(\omega)$	$Y(\omega)$
H	1	0
T	0	1

$$\{\omega \colon X(\omega) = 0, Y(\omega) = 0\} = \varnothing,$$

$$\mathrm{P}(X = 0, Y = 0) = 0$$

$$\{\omega \colon X(\omega) = 1, Y(\omega) = 0\} = \{H\},$$

$$\mathrm{P}(X = 1, Y = 0) = \frac{1}{2}$$

$X \setminus Y$	0	1
0	$p_{0,0}$	$p_{0,1}$
1	$p_{1,0}$	$p_{1,1}$

Example. Let a coin be tossed once. Let X be the number of Heads, Y the number of Tails. Then $\Omega = \{H, T\}$.

Now,

ω	$X(\omega)$	$Y(\omega)$
H	1	0
T	0	1

$$\{\omega\colon X(\omega) = 0, Y(\omega) = 0\} = \varnothing,$$

$$\mathrm{P}(X = 0, Y = 0) = 0$$

$$\{\omega\colon X(\omega) = 1, Y(\omega) = 0\} = \{H\},$$

$$\mathrm{P}(X = 1, Y = 0) = \frac{1}{2}$$

$X \setminus Y$	0	1
0	$p_{0,0}$	$p_{0,1}$
1	$p_{1,0}$	$p_{1,1}$

$X \setminus Y$	0	1
0	0	$\frac{1}{2}$
1	$\frac{1}{2}$	0

Problem. There are M white balls and $N - M$ black balls in an urn. All balls are taken out sequentially without replacement. Let X be the number of black balls taken out before the first white ball appeared, Y be the number of black balls taken out between the first white ball and the second white ball. Find $\mathrm{P}(X = k)$, $\mathrm{P}(X = k, Y = l)$.

Problem. There are M white balls and $N - M$ black balls in an urn. All balls are taken out sequentially without replacement. Let X be the number of black balls taken out before the first white ball appeared, Y be the number of black balls taken out between the first white ball and the second white ball. Find $\mathrm{P}(X = k)$, $\mathrm{P}(X = k, Y = l)$.

Solution. Let event A_i occur when the ith ball is white. Then

$$\{X = k\} = \bar{A}_1 \cap \bar{A}_2 \cap \ldots \cap \bar{A}_k \cap A_{k+1},$$

$$\mathrm{P}(X = k) = \mathrm{P}(\bar{A}_1)\mathrm{P}(\bar{A}_2|\bar{A}_1) \ldots \mathrm{P}(A_{k+1}|\bar{A}_1 \cap \ldots \bar{A}_k) =$$
$$= \frac{N-M}{N} \cdot \frac{N-M-1}{N-1} \cdots \frac{N-M-k+1}{N-k+1} \cdot \frac{M}{N-k},$$

$$\{X = k, Y = l\} = \bar{A}_1 \cap \bar{A}_2 \cap \ldots \cap \bar{A}_k \cap A_{k+1} \cap \bar{A}_{k+2} \cap \ldots \cap \bar{A}_{k+l+1} \cap A_{k+l+2},$$

$$\mathrm{P}(X = k, Y = l) = \frac{N-M}{N} \cdot \frac{N-M-1}{N-1} \cdots \frac{N-M-k+1}{N-k+1} \cdot \frac{M}{N-k} \times$$

$$\times \frac{N-M-k}{N-k-1} \cdots \frac{N-M-k-l+1}{N-k-l} \cdot \frac{M-1}{N-k-l-1} =$$

$$= \frac{(N-M)(N-M-1)\cdots(N-M-k-l+1)M(M-1)}{N(N-1)\cdots(N-k-l-1)}$$

With respect to the joint distribution the probability distribution of a single variable (e.g. X) is called *marginal distribution*. It can be easily recovered from the joint distribution:

$$
\begin{aligned}
P(X = x_i) &= P\Big(\{X = x_i\} \cap \underbrace{\bigcup_{j\geqslant 1}\{Y = y_j\}}_{=\Omega}\Big) = \\
&= P\big(\bigcup_{j\geqslant 1}\{X = x_i, Y = y_j\}\big) = \\
&= \sum_{j\geqslant 1} P(X = x_i, Y = y_j) = \sum_{j\geqslant 1} p_{i,j}, \\
P(Y = y_j) &= \sum_{i\geqslant 1} p_{i,j}.
\end{aligned}
$$

Given an event A, $\mathrm{P}(A) > 0$, the conditional distribution of random variable X can be defined as follows:

$$\mathrm{P}(X = x_i | A) = \frac{\mathrm{P}(\{X = x_i\} \cap A)}{\mathrm{P}(A)}$$

Example. (Lack-of-memory property of geometric distribution) Random variable X is said to have geometric distribution with parameter p if $\mathrm{P}(X = k) = (1-p)^{k-1}p$ for $k = 1, 2, \dots$. Then

$$\begin{aligned}\mathrm{P}(X > n) &= \mathrm{P}(X = n+1) + \mathrm{P}(X = n+2) + \dots = \\ &= (1-p)^n p + (1-p)^{n+1} p + \dots = \\ &= \frac{(1-p)^n p}{1-(1-p)} = (1-p)^n,\end{aligned}$$

$$\begin{aligned}
P(X > k+n | X > n) &= \frac{P(X > k+n, X > n)}{P(X > n)} = \\
&= \frac{P(X > k+n)}{P(X > n)} = \frac{(1-p)^{k+n}}{(1-p)^n} = (1-p)^k, \\
P(X = n+k+1 | X > n) &= P(\{X > n+k\} \setminus \{X > n+k+1\} | X > n) \\
&= (1-p)^k - (1-p)^{k+1} = (1-p)^k p.
\end{aligned}$$

As we know, X sometimes can be interpreted as expectation time for the first success in Bernoulli trials. So, the conditional distribution of expectation time given it's more that n has geometrical distribution as well.

From multiplication theorem,

$$P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j|X = x_i)$$

$$P(X = x_i, Y = y_j) = P(Y = y_j)P(X = x_i|Y = y_j)$$

given that all conditional expectations needed are defined.

Definiton

Random variables X and Y are called *independent* if

$$P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j)$$

holds for all i, j.

In other words, distribution of X does not depend on the value of Y and vice-versa.

For a fixed y_j conditional probability distribution of random variable X can be expressed in terms of joint distribution of X and Y:

$$\mathrm{P}(X = x_i | Y = y_j) = \frac{\mathrm{P}(X = x_i, Y = y_j)}{\mathrm{P}(Y = y_j)} = \\ = \frac{p_{i,j}}{\sum_{m \geqslant 1} p_{m,j}}.$$

Example. Joint distribution of random variables X, Y is given in the following table:

$X \setminus Y$	-1	0	1	
0	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{1}{16}$	
1	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{5}{16}$	

Find marginal and conditional distributions of X and Y. Check for independence.

Example. Joint distribution of random variables X, Y is given in the following table:

$X \setminus Y$	-1	0	1	marginal
0	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{1}{16}$	$\frac{1}{16}+\frac{1}{4}+\frac{1}{16}=\frac{3}{8}$
1	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{5}{16}$	$\frac{1}{16}+\frac{1}{4}+\frac{5}{16}=\frac{5}{8}$

Find marginal and conditional distributions of X and Y. Check for independence.

Example. Joint distribution of random variables X, Y is given in the following table:

$X \setminus Y$	-1	0	1	marginal
0	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{1}{16}$	$\frac{1}{16}+\frac{1}{4}+\frac{1}{16}=\frac{3}{8}$
1	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{5}{16}$	$\frac{1}{16}+\frac{1}{4}+\frac{5}{16}=\frac{5}{8}$
marginal	$\frac{2}{16}$	$\frac{2}{4}$	$\frac{6}{16}$	

Find marginal and conditional distributions of X and Y. Check for independence.

Example. Joint distribution of random variables X, Y is given in the following table:

$X \setminus Y$	-1	0	1	marginal
0	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{1}{16}$	$\frac{1}{16}+\frac{1}{4}+\frac{1}{16}=\frac{3}{8}$
1	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{5}{16}$	$\frac{1}{16}+\frac{1}{4}+\frac{5}{16}=\frac{5}{8}$
marginal	$\frac{2}{16}$	$\frac{2}{4}$	$\frac{6}{16}$	

Find marginal and conditional distributions of X and Y. Check for independence.
Since $\mathrm{P}(X=0, Y=-1)=\frac{1}{16}\neq\frac{3}{8}\cdot\frac{1}{8}=\mathrm{P}(X=0)\mathrm{P}(Y=-1)$, the two random variables are dependent.
Conditional distribution of Y:

$X \setminus Y$	-1	0	1	marginal
0	$\frac{1\cdot 8}{16\cdot 3}=\frac{1}{6}$	$\frac{1\cdot 8}{4\cdot 3}=\frac{2}{3}$	$\frac{1\cdot 8}{16\cdot 3}=\frac{1}{6}$	$\frac{3}{8}$
1	$\frac{1\cdot 8}{16\cdot 5}=\frac{1}{10}$	$\frac{1\cdot 8}{4\cdot 5}=\frac{2}{5}$	$\frac{5\cdot 8}{16\cdot 5}=\frac{1}{2}$	$\frac{5}{8}$
marginal	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{3}{8}$	

Theorem

Let X and Y be independent nonegative integer random variables. Then

$$P(X+Y=m)=\sum_{k=0}^{m} P(X=k)P(Y=m-k), \qquad m=0,1,\dots$$

Proof.

$$\begin{aligned}
P(X+Y=m) &= P\Big(\{X+Y=m\}\cap\bigcup_{k=0}^{\infty}\{X=k\}\Big)\\
&= \sum_{k=0}^{\infty} P(X+Y=m, X=k) = \sum_{k=0}^{\infty} P(k+Y=m, X=k)\\
&= \sum_{k=0}^{\infty} P(Y=m-k, X=k) = \sum_{k=0}^{\infty} P(X=k)P(Y=m-k)\\
&= \sum_{k=0}^{m} P(X=k)P(Y=m-k).
\end{aligned}$$

Sum of independent Poissonian random variables. Let X have Poisson distribution with parameter λ, Y Poisson distribution with parameter μ. Then according to previous theorem,

$$\begin{aligned} \mathrm{P}(X+Y=m) &= \sum_{k=0}^{m} \frac{\lambda^k}{k!} e^{-\lambda} \cdot \frac{\mu^{m-k}}{(m-k)!} e^{-\mu} = \\ &= e^{-(\lambda+\mu)} \frac{1}{m!} \sum_{k=0}^{m} \frac{m!}{k!(m-k)!} \lambda^k \mu^{m-k} \\ &= \frac{(\lambda+\mu)^m}{m!} e^{-(\lambda+\mu)}. \end{aligned}$$

In othe words, class of poissonian random variables is closed under summation.

Let X and Y be independent integer random variables. As we have just seen, the joint distribution of a term and a sum is given by the following formula:

$$\mathrm{P}(X = k, X + Y = m) = \mathrm{P}(X = k)\mathrm{P}(Y = m - k).$$

Applying it to the previous example, obtain:

$$\begin{aligned}\mathrm{P}(X = k | X + Y = m) &= \frac{\frac{\lambda^k}{k!}e^{-\lambda} \cdot \frac{\mu^{m-k}}{(m-k)!}e^{-\mu}}{\frac{(\lambda+\mu)^m}{m!}e^{-(\lambda+\mu)}} \\ &= \frac{m!}{k!(m-k)!}\left(\frac{\lambda}{\lambda+\mu}\right)^k\left(\frac{\mu}{\lambda+\mu}\right)^{m-k}\end{aligned}$$

The conditional distribution of one term is binomial. Nota bene: sometimes known distributions appear not from their usual physical origins!

Problem. Ann and Bart roll one die each untill ⚅ appears. The one who gets ⚅ first recieves one rouble from other and the game is over. If they both have ⚅, it's a tie and nobody pays. Let X denote the number of rolls until the end of game, Y denote Ann's gain. Find joint distribution of X and Y and then find marginal probability distributions of X and Y.

Problem. Ann and Bart roll one die each untill ⚅ appears. The one who gets ⚅ first recieves one rouble from other and the game is over. If they both have ⚅, it's a tie and nobody pays. Let X denote the number of rolls until the end of game, Y denote Ann's gain. Find joint distribution of X and Y and then find marginal probability distributions of X and Y.

Solution. Y is either -1 rouble, 0 roubles, or 1 rouble. Then event $\{X = k, Y = -1\}$ occurs if there were $k - 1$ non-⚅ on both dice, and for last time Bart had ⚅ while Ann hadn't. Since all trials are independent,

$$\mathrm{P}(X = k, Y = -1) = \left(\frac{5}{6} \cdot \frac{5}{6}\right)^{k-1} \cdot \frac{5}{6} \cdot \frac{1}{6} = \left(\frac{5}{6}\right)^{2k-1} \frac{1}{6}.$$

$$P(X = k, Y = 1) = \left(\frac{5}{6} \cdot \frac{5}{6}\right)^{k-1} \cdot \frac{1}{6} \cdot \frac{5}{6} = \left(\frac{5}{6}\right)^{2k-1} \frac{1}{6},$$

$$P(X = k, Y = 0) = \left(\frac{5}{6} \cdot \frac{5}{6}\right)^{k-1} \cdot \frac{1}{6} \cdot \frac{1}{6} = \left(\frac{5}{6}\right)^{2k-2} \frac{1}{36},$$

Now,

$$\begin{aligned} P(X = k) &= P(X = k, Y = -1) + P(X = k, Y = 0) + P(X = k, Y = 1) = \\ &= \left(\frac{5}{6}\right)^{2k-2} \left(\frac{5}{36} + \frac{1}{36} + \frac{5}{36}\right) = \\ &= \left(\frac{25}{36}\right)^{k-1} \frac{11}{36}. \end{aligned}$$

X has geometric distribution with parameter $\frac{11}{36}$ (non wonder!).

$$
\begin{aligned}
P(Y=-1) &= \sum_{k=1}^{\infty} P(X=k, Y=-1) = \\
&= \sum_{k=1}^{\infty} \left(\frac{5}{6}\right)^{2k-1} \frac{1}{6} = \frac{5}{6} \cdot \frac{1}{6} \frac{1}{1-\frac{25}{36}} = \frac{5}{11}, \\
P(Y=1) &= \ldots = \frac{5}{11}, \\
P(Y=0) &= \sum_{k=1}^{\infty} P(X=k, Y=0) = \\
&= \sum_{k=1}^{\infty} \left(\frac{5}{6}\right)^{2k-2} \frac{1}{36} = \frac{1}{36} \frac{1}{1-\frac{25}{36}} = \frac{1}{11}.
\end{aligned}
$$

$$\mathrm{P}(X=k, Y=-1)=\left(\frac{5}{6}\right)^{2k-1}\frac{1}{6}=\left(\frac{5}{6}\right)^{2k-2}\frac{11}{36}\cdot\frac{5}{11}=\mathrm{P}(X=k)\mathrm{P}(Y=-1).$$

$$\mathrm{P}(X=k, Y=1)=\left(\frac{5}{6}\right)^{2k-1}\frac{1}{6}=\left(\frac{5}{6}\right)^{2k-2}\frac{11}{36}\cdot\frac{5}{11}=\mathrm{P}(X=k)\mathrm{P}(Y=1).$$

$$\mathrm{P}(X=k, Y=0)=\left(\frac{5}{6}\right)^{2k-2}\frac{1}{36}=\left(\frac{5}{6}\right)^{2k-2}\frac{11}{36}\cdot\frac{1}{11}=\mathrm{P}(X=k)\mathrm{P}(Y=0).$$

We see that random variables X and Y are independent.