APPENDIX
to the article by M.I.Schlesinger

MAXIMUM LIKELIHOOD ESTIMATION OF THE CERTAIN MODEL OF CONDITIONAL INDEPENDENCE

(for the reviewing process)

Proof of the theorem 1. We will prove, that, if such community $p^*(x,y,k),\ x\in X,\ y\in Y,\ k\in K$, exists, that satisfies the conditions (7-12), then using the numbers $p^*(x,y),\ x\in X,\ y\in Y$, a single pair of functions may be restored, one of them (unknown which) being $p^*(y\ /\ k=1)$, and second being $p^*(y\ /\ k=2)$.

Let us represent the numbers $p^*(x,y),\ x\in X,\ y\in Y$, by the table, that consists of $|X|$ columns and $|Y|$ rows. The rows of the table correspond to the values of feature y , columns corresponding to values of the feature x , the value $p^*(x,y)$ being written on the crossing of x -th column and y -th row.

Let us consider the columns of the table as some $|Y|$ - dimensional vectors and let us build a linear closure of this set of vectors. This linear closure forms two-dimensional linear space. Really, for any $x\in X$ vector in x -th column, i.e. the vector with components $p^*(x,y), y\in Y$, is a linear combination of two $|Y|$ -dimensional vectors whose components are $p^*(y\ /\ k=1),\ y\in Y$, and $p^*(y\ /\ k=2),\ y\in Y$, correspondingly, because due to the conditions (7) and (12) for every $x\in X$ holds

$$p^*(x,y)=p^*(k=1,x)\cdot p^*(y\ /\ k=1)+p^*(k=2,x)\cdot p^*(y\ /\ k=2)\ . \tag{25}$$

As to vectors $p^*(y\ /\ k=1)$ and $p^*(y\ /\ k=2)$, they are not linearly dependent, because they are not null-vectors and due to the condition (10) they are not equal one to another and, consequently, they are not collinear one with another.

The set of columns of the table contains in itself the columns, which correspond to the ideal representatives, whose existence is guaranteed due to the conditions (8) and (9). These columns may be detected by the following rules.

1. Let us exclude from the table the columns with null-vectors. Every such excluded column corresponds to some value of x , for which $p^*(x)=0$. Such column does not correspond to an ideal representative. Really, a probability $p^*(x)$ for ideal representative cannot be zero, because $p^*(x)=p^*(k=1)\cdot p^*(x\ /\ k=1)+p^*(k=2)\cdot p^*(x\ /\ k=2)$. In this expression the probabilities $p^*(k=1)$ and $p^*(k=2)$ are not zero due to the condition (11), and for the ideal representative one of the probabilities $p^*(x/\,k=1)$ or $p^*(x/\,k=2)$ should not be equal to zero due to the conditions (8) and (9).

2. Let us exclude from the rest of the table such columns, which correspond to those values x , for which such values x_1 and x_2 and such numbers $\alpha_1 > 0$ and $\alpha_2 > 0$ exist, that equality

$$p^*(x,y) = \alpha_1 \cdot p^*(x_1,y) + \alpha_2 \cdot p^*(x_2,y) \tag{26}$$

holds for every $y \in Y$, and equality

$$p^*(x_1,y) = p^*(x_2,y) \tag{27}$$

does not hold for some $y \in Y$. Let us prove, that such columns do not correspond to the ideal representatives. Taking on account (25) the condition (26) may be transformed into the form

$$\begin{aligned} p^*(x,y) &= \alpha_1 \cdot p^*(x_1,y) + \alpha_2 \cdot p^*(x_2,y) = \\ &= \left(\alpha_1 \cdot p^*(k=1,x_1) + \alpha_2 \cdot p^*(k=1,x_2)\right) \cdot p^*(y \,/\, k=1) + \\ &+ \left(\alpha_1 \cdot p^*(k=2,x_1) + \alpha_2 \cdot p^*(k=2,x_2)\right) \cdot p^*(y \,/\, k=2) . \end{aligned}$$

This equality implies, that

$$p^*(k=1,x) = \alpha_1 \cdot p^*(k=1,x_1) + \alpha_2 \cdot p^*(k=1,x_2) \tag{28}$$

and

$$p^*(k=2,x) = \alpha_1 \cdot p^*(k=2,x_1) + \alpha_2 \cdot p^*(k=2,x_2) \ . \tag{29}$$

It means in its turn, that $p^*(k=1,x) \neq 0$ and $p^*(k=2,x) \neq 0$, because $\alpha_1 > 0$, $\alpha_2 > 0$ and the probabilities $p^*(k=1,x_1)$ and $p^*(k=1,x_2)$ cannot both be equal zero, because due to (27) they are not equal one with another. But for ideal representative one of the probabilities $p^*(k=1,x)$ and $p^*(k=2,x)$ must be equal zero, so the value x is not ideal representative.

After such exclusions only those columns remain in the table, which correspond to those values x , for which either $p^*(k=1,x)=0$ and $p^*(k=2,x)\neq 0$, or $p^*(k=1,x)\neq 0$ and $p^*(k=2,x)=0$, i.e. which correspond to the ideal representatives. In any column in this rest of the table the vector with such components $p^*(x,y)$, $y \in Y$, is written, which are equal either to $p^*(k=1,x)\cdot p^*(y \,/\, k=1)$, $y \in Y$, or $p^*(k=2,x)\cdot p^*(y \,/\, k=2)$, $y \in Y$. Consequently, a function $\dfrac{p^*(x,y)}{\sum_{y \in Y} p^*(x,y)}$, which is to be calculated for every not excluded x -th column, equals either $p^*(y \,/\, k=1)$ or $p^*(y \,/\, k=2)$.

Using the obtained probabilities $p^*(x,y)$ as the base the unique pair of functions may be built, one of them (unknown which) being $p^*(y \,/\, k=1)$ and another being $p^*(y \,/\, k=2)$. It means, that the function $p^*(y\ /\ k)$ may be restored up to a permutation of values of the state k. For every of these two variants for $p^*(y\ /\ k)$ the function $p^*(x,k)$ also may be uniquely restored. It follows from the fact, that for every value x the function $p^*(x,k)$ may be uniquely represented in the form (25). As to probabilities $p^*(x,y,k)$, they are to be calculated by simple multiplication of the probabilities $p^*(x,k)$ and $p^*(y/k)$. **The theorem is proved.**

Proof of the theorem 2. Let us exclude the auxiliary variables $\alpha_k(x,y)$ from the equations (17-20). To make that one must substitute the expression (17) for $\alpha_k(x,y)$ into (18), (19), (20). As a result the following equations are obtained:

$$p^{t+1}(k) = \sum_{x,y} \frac{p^*(x,y)}{p^t(x,y)} \cdot p^t(k) \cdot p^t(x\ /\ k) \cdot p^t(y\ /\ k), \quad k \in K\ ; \tag{30}$$

$$p^{t+1}(x\ /\ k) = \frac{\sum\limits_{y} \frac{p^*(x,y)}{p(x,y)} \cdot p^t(k) \cdot p^t(x\ /\ k) \cdot p^t(y\ /\ k)}{p^{t+1}(k)}, \quad k \in K,\ x \in X\ ; \tag{31}$$

$$p^{t+1}(y\ /\ k) = \frac{\sum\limits_{x} \frac{p^*(x,y)}{p(x,y)} \cdot p^t(k) \cdot p^t(x\ /\ k) \cdot p^t(y\ /\ k)}{p^{t+1}(k)}, \quad k \in K,\ y \in Y\ . \tag{32}$$

A condition of the point stability means, that the values $p^t(k), p^t(x\ /\ k)$ and $p^t(y\ /\ k)$ in the expressions (30), (31), (32) are equal to the values $p^{t+1}(k), p^{t+1}(x\ /\ k), p^{t+1}(y\ /\ k)$ correspondingly, i.e. one can simply omit the indices t and $t+1$ in these expressions. After such transformations the conditions (30), (31), (32) may be rewritten in the form

$$\sum_{x,y} \frac{p^*(x,y)}{p(x,y)} \cdot p(x\ /\ k) \cdot p(y\ /\ k) = 1, \ \ k \in K\ ; \tag{33}$$

$$\sum_{y} \frac{p^*(x,y)}{p(x,y)} \cdot p(y\ /\ k) = 1, \ \ k \in K,\ x \in X\ ; \tag{34}$$

$$\sum_{x} \frac{p^*(x,y)}{p(x,y)} \cdot p(x\ /\ k) = 1, \ \ k \in K,\ y \in Y\ . \tag{35}$$

Let us clarify, what values shall take the coefficients λ, η_k and γ_k in the equations (21,a,b,c) taking on account, that the conditions (21, d,e,f) must be satisfied. Let us multiply the equation (21,a) by $p(k)$, calculate the sum of obtained products over all k. As a result, we will receive, that

$$\sum_k \lambda \cdot p(k) = \lambda = -\sum_k p(k) \sum_{x,y} \frac{p^*(x,y)}{p(x,y)} \cdot p(x/k) \cdot p(y/k) =$$

$$= -\sum_{x,y} \frac{p^*(x,y)}{p(x,y)} \sum_k p(k) \cdot p(x/k) \cdot p(y/k) = -\sum_{x,y} p^*(x,y) = -1 \ ,$$

i.e. λ =-1 ; so the condition (21) may be performed in the form

$$\sum_{x,y} \frac{p^*(x,y)}{p(x,y)} \cdot p(x/k) \cdot p(y/k) = 1 \ , \qquad (36)$$

that coincides exactly with (33).

Let us multiply (21,b) by $p(x/k)$, calculate the sum of these products over all x and as a result we will receive that

$$\sum_x \gamma_k \cdot p(x/k) = \gamma_k = -\sum_x p(x/k) \sum_y \frac{p^*(x,y)}{p(x,y)} \cdot p(k) \cdot p(y/k) =$$

$$= -p(k) \sum_{x,y} \frac{p^*(x,y)}{p(x,y)} \cdot p(x/k) \cdot p(y/k) \ .$$

Taking on account (36) one can see, that $\gamma_k = -p(k)$ and that the condition (21,b) occurs to be equivalent to the condition (34).

Similarly, after multiplication of equation (21,c) by $p(y/k)$ and summation the obtained products over all values of y it becomes clear, that $\eta_k = -p(k)$ and that the condition (21,ñ) is the same as (35). **The theorem is proved.**

Proof of the theorem 3. Naturally, to prove the theorem it is necessary only to prove, that if at least one of the inequalities $p(x/k) \neq p(x)$ or $p(y/k) \neq p(y)$ fulfills, then the conditions of the theorem imply the inequality (24). We consider the proof for the case, when $p(y/k) \neq p(y)$. This proof is given by the following 13 proved statements.

1. The equality (22) implies the following equalities:

$$\sum_k p(k) \cdot p(y/k) \sum_x \frac{p^*(x,y) \cdot p(x/k)}{\sum_k p(k) \cdot p(x/k) \cdot p(y/k)} = \sum_k p(k) \cdot p(y/k) \ ;$$

$$\sum_x \frac{p^*(x,y) \cdot \sum_k p(k) \cdot p(x/k) \cdot p(y/k)}{\sum_k p(k) \cdot p(x/k) \cdot p(y/k)} = \sum_k p(y,k) \ ;$$

$$\sum_x p^*(x,y) = p(y) \ ;$$

$$p^*(y) = p(y) \ . \tag{37}$$

Similarly one can prove, that the equality (23) implies the equality

$$p^*(x) = p(x) \ . \tag{38}$$

2. Due to (37) the equality (22) may be written in the form

$$\sum_x \frac{p^*(x \ / \ y)}{p(x \ / \ y)} \cdot p(x \ / \ k) = 1 \quad , \tag{39}$$

and the equality (23) due to (38) in the form

$$\sum_y \frac{p^*(y \ / \ x)}{p(y \ / \ x)} \cdot p(y \ / \ k) = 1, \ \ \forall k, \ \forall x \ . \tag{40}$$

3. Let x' be some value of variable x ; then for every value x such number $\alpha(x)$ exists, that the equality

$$p^*(y \ / \ x') \text{-} \ p^*(y) = \alpha(x) \cdot \ \left(p^*(y \ / \ x) - p^*(y)\right) \tag{41}$$

is fulfilled for every value y , provided at least for some value of y the inequality $p^*(y \ / \ x) \neq p^*(y)$ holds. Really, the following inequalities are valid:

$$p^*(y) = p^*(k=1) \cdot p^*(y \ / \ k=1) + p^*(k=2) \cdot p^*(y \ / \ k=2) \ ; \tag{42}$$

$$p^*(y \ / \ x) = p^*(k=1 \ / \ x) \cdot p^*(y \ / \ k=1) + p^*(k=2 \ / \ x) \cdot p^*(y \ / \ k=2) \ ; \tag{43}$$

$$p^*(y \ / \ x') = p^*(k=1 \ / \ x') \cdot p^*(y \ / \ k=1) + p^*(k=2 \ / \ x') \cdot p^*(y \ / \ k=2) \ . \tag{44}$$

The equalities(42) and (43) imply the equality

$$p^*(y \ / \ x) - p^*(y) = \left[p^*(k=1 \ / \ x) - p^*(k=1)\right] \cdot \left[p^*(y \ / \ k=1) - p^*(y \ / \ k=2)\right] \ , \tag{45}$$

and the equalities (42) and (44) imply the equality

$$p^*(y \ / \ x') - p^*(y) = \left[p^*(k=1 \ / \ x') - p^*(k=1)\right] \cdot \left[p^*(y \ / \ k=1) - p^*(y \ / \ k=2)\right] \ . \tag{46}$$

The inequality $p^*(k=1 \ / \ x) - p^*(k=1) \neq 0$ is valid , because the inequality $p^*(y \ / \ x) \neq p^*(y)$ is fulfilled at least for some values of y , and so the expression

$$p^*(y \ / \ x') \text{-} \ p^*(y) = \frac{p^*(k=1 \ / \ x') - p^*(k=1)}{p^*(k=1 \ / \ x) - p^*(k)} \left[p^*(y \ / \ x) - p^*(y)\right] \ , \tag{47}$$

may be obtained from (45) and (46) and so the statement (41) is proved.

4. Lemma 1. Let x_1 and x_2 be such two values of the variable x , that the numbers $q_1 = p^*(k=1 \ / \ x_1) - p^*(k=1)$ and $q_2 = p^*(k=1 \ / \ x_2) - p^*(k=1)$ have opposite signs in the sense, that $q_1 \cdot q_2 \leq 0$. Then the numbers

$r_1 = p(k=1\ / x_1) - p(k=1)$ and $r_2 = p(k=1\ / x_2) - p(k=1)$ have also opposite signs in the sense, that $r_1 \cdot r_2 \le 0$.

Proof. As far the numbers q_1 and q_2 have opposite signs, the probability $p^*(k=1)$ is a convex linear combination of probabilities $p^*(k=1\ / x_1)$ and $p^*(k=1\ / x_2)$ and such number β, $0 \le \beta \le 1$, exists, that for every y

$$p^*(y) = \beta \cdot p^*(y\ / x_1) + (1-\beta) \cdot p^*(y\ / x_2) \ . \tag{48}$$

The equality

$$\sum_y \frac{p^*(y\ / x)}{p(y\ / x)} \cdot \left[p(y\ / k=1) - p(y\ / k=2)\right] = 0 \tag{49}$$

is valid for every x , because the equality (40) holds for every x and every k . Consequently, (49) holds also for x_1 and x_2 , i.e.

$$\sum_y \frac{p^*(y\ / x_1)}{p(y\ / x_1)} \cdot \left[p(y\ / k=1) - p(y\ / k=2)\right] = 0 \ , \tag{50}$$

$$\sum_y \frac{p^*(y\ / x_2)}{p(y\ / x_2)} \cdot \left[p(y\ / k=1) - p(y\ / k=2)\right] = 0. \tag{51}$$

It is clear also, that

$$\sum_y \frac{p^*(y)}{p(y)} \cdot \left[p(y\ / k=1) - p(y\ / k=2)\right] = 0 \ , \tag{52}$$

because it was proved before (see statement 1), that $p^*(y) = p(y)$.

Let us assume for certainty sake, that $p(k=1\ / x_1) < p(k=1\ / x_2)$, and prove, that in this case the probability $p(k=1)$ lies in the range from $p(k=1\ / x_1)$ to $p(k=1\ / x_2)$ and so the numbers $p(k=1\ / x_1) - p(k=1)$ and $p(k=1\ / x_2) - p(k=1)$ have opposite signs.

Let us consider the following two expressions, which depend on some variable coefficient γ :

$$\sum_y \frac{p^*(y\ / x_1) \cdot \left[p(y\ / k=1) - p(y\ / k=2)\right]}{p(y\ / k=2) + \gamma \cdot \left[p(y\ / k=1) - p(y\ / k=2)\right]} \tag{53}$$

and

$$\sum_y \frac{p^*(y\ / x_2) \cdot \left[p(y\ / k=1) - p(y\ / k=2)\right]}{p(y\ / k=2) + \gamma \cdot \left[p(y\ / k=1) - p(y\ / k=2)\right]} \ . \tag{54}$$

Both a value (53), and a value (54) strictly decrease, when coefficient γ increases, because at least for some values of y the inequality $p(y / k=1) \ne p(y / k=2)$ holds.

Remark. The monotonous dependence of the values (53) and (54) on the value γ implies the following fact, which will be used later. Namely, if for some x the equality $p^*(y/x) = p^*(y)$ holds, then for the same value x the similar equality $p(y/x) = p(y)$ holds too. Really, the condition $\sum_y \frac{p^*(y)}{p(y/x)} \cdot p(y/k) = 1,\ k = 1{,}2$, implies the equality $\sum_y \frac{p^*(y) \cdot [p(y/k=1) - p(y/k=2)]}{p(y/k=2) + p(k=1/x) \cdot [p(y/k=1) - p(y/k=2)]} = 0$, that can be true if and due to the monotonous dependence only if $p(k=1/x) = p(k=1)$, and, consequently, it and only if $p(y/x) = p(y)$. E**nd of the remark.**

The value (53) equals (50) when $\gamma = p(k=1/x_1)$ and the value (54) equals (51) when $\gamma = p(k=1/x_2)$. Consequently, taking on account, that the value (53) depends on γ strictly monotonously, the inequalities

$$\sum_y \frac{p^*(y/x_1) \cdot [p(y/k=1) - p(y/k=2)]}{p(y/k=2) + \gamma \cdot [p(y/k=1) - p(y/k=2)]} > 0 \ , \qquad (55)$$

$$\sum_y \frac{p^*(y/x_2) \cdot [p(y/k=1) - p(y/k=2)]}{p(y/k=2) + \gamma \cdot [p(y/k=1) - p(y/k=2)]} > 0 \ , \qquad (56)$$

are valid for every $\gamma < p(k=1/x_1)$, and the inequalities

$$\sum_y \frac{p^*(y/x_1) \cdot [p(y/k=1) - p(y/k=2)]}{p(y/k=2) + \gamma \cdot [p(y/k=1) - p(y/k=2)]} < 0 \ , \qquad (57)$$

$$\sum_y \frac{p^*(y/x_2) \cdot [p(y/k=1) - p(y/k=2)]}{p(y/k=2) + \gamma \cdot [p(y/k=1) - p(y/k=2)]} < 0 \qquad (58)$$

are valid for every $\gamma > p(k=2/x_1)$. We will add inequalities (55) and (56) with the weights β and $1-\beta$, and taking on account (48) we will receive, that for every $\gamma < p(k=1/x_1)$

$$\sum_y \frac{p^*(y) \cdot [p(y/k=1) - p(y/k=2)]}{p(y/k=2) + \gamma \cdot [p(y/k=1) - p(y/k=2)]} > 0 \ , \qquad (59)$$

and for every $\gamma > p(k=1/x_2)$

$$\sum_y \frac{p^*(y) \cdot [p(y/k=1) - p(y/k=2)]}{p(y/k=2) + \gamma \cdot [p(y/k=1) - p(y/k=2)]} < 0 \ . \qquad (60)$$

When $\gamma = p(k=1)$ the values on the left sides in (59) and (60) become equal to the left sides of (52) and due to (52) become equal to zero. Consequently $p(k=1)$ lies obligatory in the range from $p(k=1/x_1)$ to $p(k=1/x_2)$, because in any other cases the left sides in (59) and (60) do not equal to zero. So the differences $p(k=1/x_1)-p(k=1)$ and $p(k=1/x_2)-p(k=1)$ have different signs. **The lemma is proved.**

5. Lemma 2. For every x the difference $p(k=1/x)-p(k=1)$ and the difference $p(x/k=1)-p(x/k=2)$ have the same signs.

Proof. Let $p(x/k=1)=p(x/k=2)$. In this case

$$p(k=1/x)=\frac{p(k=1)\cdot p(x/k=1)}{p(k=1)\cdot p(x/k=1)+p(k=2)\cdot p(x/k=2)}=\frac{p(k=1)}{p(k=1)+p(k=2)}=p(k=1) .$$

Let $p(x/k=1)>p(x/k=2)$ and let us denote $\gamma=\frac{p(x/k=1)}{p(x/k=2)}>1$. Then one can receive, that

$$p(k=1/x)=p(k=1)\cdot\frac{1}{p(k=1)+\frac{p(k=2)}{\gamma}}>p(k=1) .$$ **The lemma is proved.**

6. Corollary. Let x^* be some value of variable x , such that $p^*\left(k=1/x^*\right)\neq p^*(k=1)$. Then the value

$$\left[p(x/k=1)-p(x/k=2)\right]\frac{p^*(k=1/x)-p^*(k=1)}{p^*\left(k=1/x^*\right)-p^*(k=1)} \tag{61}$$

is either non-positive for every $x\in X$ or non-negative for every $x\in X$.

Proof. Really, the denominator in the expression (61) does not depend on x . So, if the value of (61) would change its sign under changes of x , one and only one of the values of $p(x/k=1)-p(x/k=2)$ or $p^*(k=1/x)-p^*(k=1)$ would change its signs. It would contradict either the lemma 1 or the lemma 2. **The corollary is proved.**

7. While the equality (39) is valid both for $k=1$ and for $k=2$, it implies the equalities

$$\sum_x\frac{p^*(x/y)}{p(x/y)}\left[p(x/k=1)-p(x/k=2)\right]=0 , \tag{62}$$

$$\sum_x\frac{p^*(x/y)}{p(x/y)}p(x)=1 . \tag{63}$$

8. Let x^* be such value of x , that at least for some values of y the inequality $p^*\left(y/x^*\right)\neq p^*(y)$ is fulfilled. The further considerations are to be fulfilled separately for the

following two cases: when such x^* exists and when it does not exist. Let us assume at first, that such value x^* exists and this value will be fixed during the further considerations. The case, when such value does not exists, will be discussed later, in the statement 12.

Because (62) holds for every value y , the equation

$$\sum_y \left[p^*\left(y/x^*\right)-p^*(y)\right]\sum_x \frac{p^*(x/y)}{p(x/y)}\left[p(x/k=1)-p(x/k=2)\right]=0 ,$$

is valid too. It is not difficult to transform it into the form

$$\sum_x \left[p(x/k=1)-p(x/k=2)\right]\sum_y \frac{p^*(y/x)}{p(y/x)}\left[p^*\left(y/x^*\right)-p^*(y)\right]=0 . \tag{64}$$

This transformation is based on the validity of the following sequence of the equalities, $\frac{p^*(x/y)}{p(x/y)}=\frac{p^*(y)\cdot p^*(y/x)}{p(y)\cdot p(y/x)}=\frac{p^*(x)\cdot p^*(y/x)}{p(x)\cdot p(y/x)}=\frac{p^*(y/x)}{p(y/x)}$, which are valid because $p^*(x)=p(x)$ and $p^*(y)=p(y)$.

9. Let us consider the sum $f(x)=\sum_y \frac{p^*(y/x)}{p(y/x)}\left[p^*\left(y/x^*\right)-p^*(y)\right]$. We will divide the set X onto two subsets: $X^=$ and $X^{\neq}$. The first of them contains all those values of x , for which $p^*(y/x)=p^*(y)$, and, consequently, $p(y/x)=p(y)$ holds. The second one, i.e. the subset $X^{\neq}$, contains all other values of x .

For every $x\in X^=$ holds

$$f(x)=\sum_y \frac{p^*(y)}{p(y)}\left[p^*\left(y/x^*\right)-p^*(y)\right]=\sum_y p^*\left(y/x^*\right) - \sum_y p^*(y)=0 .$$

For all other values of x an equation

$$f(x)=\frac{p^*\left(k=1/x^*\right)-p^*(k=1)}{p^*(k=1/x)-p^*(k)}\sum_y \frac{p^*(y/x)}{p(y/x)}\left[p^*(y/x)-p^*(y)\right] .$$

is valid due to (47). A substitution of this expression in (64) gives

$$\sum_{x\in X^{\neq}}\left[p(x/k=1)-p(x/k=2)\right]\frac{p^*\left(k=1/x^*\right)-p^*(k=1)}{p^*(k=1/x)-p^*(k=1)}\sum_y \frac{p^*(y/x)}{p(y/x)}\left[p^*(y/x)-p^*(y)\right]=0 \tag{65}$$

10. The sum $g(x)=\sum_y \frac{p^*(y/x)}{p(y/x)}\left[p^*(y/x)-p^*(y)\right]$, that comes into the expression (65) possesses a following property.

Lemma 3. The inequality

$$g(x)=\sum_y \frac{p^*(y/x)}{p(y/x)}\left[p^*(y/x)-p^*(y)\right]\geq 0 \ ,$$

is valid for every x, the equality holding if and only if for every value of y an equality $p^*(y/x)=p(y/x)$ is true.

Proof. Taking on account, that $p^*(y)=p(y)$, one can rewrite the sum under analysis in the form

$$g(x)=\sum_y \frac{p^*(y/x)}{p(y/x)}p^*(y/x)-\sum_y \frac{p^*(y/x)}{p(y/x)}p(y) \ .$$

It has been proved before , that

$$\sum_y \frac{p^*(y/x)}{p(y/x)}p(y)=1 \ . \tag{66}$$

Let us prove, that

$$\sum_y \frac{p^*(y/x)}{p(y/x)}p^*(y/x)\geq 1 \ , \tag{67}$$

and clarify the condition, under which the inequality (67) becomes an equality. To do it one must find a function $p(y/x)$, which minimizes a left side of (67) under a restriction $\sum_y p(y/x)=1$. A function by Lagrange, that corresponds to such optimization problem, is $H=\sum_y \frac{p^*(y/x)}{p(y/x)}p^*(y/x)+\lambda\sum_y p(y/x)$, and a conditioned minimum under the search is a solution of the following system of equations

$$\frac{\partial H}{\partial p(y/x)}=-\frac{p^*(y/x)\cdot p^*(y/x)}{p^2(y/x)}+\lambda=0, \ y\in Y \ ,$$

which has a single solution $p(y/x)=p^*(y/x)$. A substitution of this solution into the right side of (67) shows, that minimal value of the right side is 1. So the inequality (67) is proved, as well as the lemma at all. **The lemma is proved.**

11. The equality (65) may be written shortly as

$$\sum_{x\in X^{\neq}} h(x)\cdot g(x)=0 \ , \tag{68}$$

with

$$h(x)=\left[p(x/k=1)-p(x/k=2)\right]\frac{p^*\left(k=1/x^*\right)-p^*(k=1)}{p^*(k=1/x)-p^*(k=1)} \ .$$

Due to the corollary (see statement 6) all numbers $h(x)$ in the expression (68) have the same sign: either all of them are positive or all of them are negative and no of them are zero, because the set $X^{\neq}$ is so arranged. As it has been just now proved (see lemma 3) a function

$g(x)$ takes only non-negative values. Consequently, the sum $\sum_{x\in X^{\neq}} h(x)\cdot g(x)$ can be zero only if every component in this sum is zero too. It means, that $\sum_{y}\frac{p^*(y/x)}{p(y/x)}\left[p^*(y/x)-p^*(y)\right]=0$ for every $x\in X^{\neq}$, but in its turn it is possible only if $p(y/x)=p^*(y/x)$ for every $y\in Y, x\in X^{\neq}$. As to the values $x\in X^{=}$, it is true by definition, that $p^*(y/x)=p^*(y)$, and so $p(y/x)=p(y)$. Consequently, the equality $p^*(y/x)=p(y/x)$ holds for $x\in X^{=}$ too.

12. So it is already proved, that the equality

$$p^*(y/x)=p(y/x) \tag{69}$$

holds for every $y\in Y$ and $x\in X$, provided such value x^* exists, that at least for some values of y the inequality $p^*\left(y/x^*\right)\neq p^*(y)$ is fulfilled. Let us consider now the case, when such value x^* does not exist. It means, that for every x and y the equality $p^*(y/x)=p^*(y)$ is fulfilled. It implies, that for every x and y also an equality $p^*(x/y)=p^*(x)$ takes place. In this case the conditions of the theorem get a form

$$\sum_{x}\frac{p^*(x)}{p(x/y)}p(x/k)=1,\ \sum_{y}\frac{p^*(y)}{p(y/x)}p(y/k)=1\ .$$

Because these equalities hold both for $k=1$ and for $k=2$, they imply the equalities $\sum_{x}\frac{p^*(x)}{p(x/y)}p(x)=1$ and $\sum_{y}\frac{p^*(y)}{p(y/x)}p(y)=1$. Taking on account, that $p(x)=p^*(x)$, and $p(y)=p^*(y)$, one can obtain the equalities $\sum_{y}\frac{p^*(y)\cdot p^*(y)}{p(y/x)}=1$ and $\sum_{x}\frac{p^*(x)\cdot p^*(x)}{p(x/y)}=1$, which due to the lemma 3 may be fulfilled only if the equations $p(y/x)=p^*(y)$ and $p(x/y)=p^*(x)$ are valid for every x and y .

13. So the equality $p(y/x)=p^*(y/x)$ holds for every x and y . It had been shown before, that equality $p(x)=p^*(x)$ holds for every x . So the equality $\sum_{k}p(k)\cdot p(x/k)\cdot p(y/x)=p^*(x,y)$ holds for every x and y . It implies quite clearly the inequality (36), which was to be proved. **The theorem is proved.**