

Conditional expectation

In probability theory, the **conditional expectation**, **conditional expected value**, or **conditional mean** of a random variable is its expected value – the value it would take “on average” over an arbitrarily large number of occurrences – given that a certain set of "conditions" is known to occur. If the random variable can take on only a finite number of values, the “conditions” are that the variable can only take on a subset of those values. More formally, in the case when the random variable is defined over a discrete probability space, the "conditions" are a partition of this probability space.

With multiple random variables, for one random variable to be mean independent of all others both individually and collectively means that each conditional expectation equals the random variable's (unconditional) expected value. This always holds if the variables are independent, but mean independence is a weaker condition.

Depending on the nature of the conditioning, the conditional expectation can be either a random variable itself or a fixed value. With two random variables, if the expectation of a random variable \mathbf{X} is expressed conditional on another random variable \mathbf{Y} without a particular value of \mathbf{Y} being specified, then the expectation of \mathbf{X} conditional on \mathbf{Y} , denoted $E(\mathbf{X}|\mathbf{Y})$, is a function of the random variable \mathbf{Y} and hence is itself a random variable. Alternatively, if the expectation of \mathbf{X} is expressed conditional on the occurrence of a particular value of \mathbf{Y} , denoted \mathbf{y} , then the conditional expectation $E(\mathbf{X}|\mathbf{Y} = \mathbf{y})$ is a fixed value.

This concept generalizes to any probability space using measure theory.

In modern probability theory the concept of conditional probability is defined in terms of conditional expectation.

Contents

Examples

History

Classical definition

- Conditional expectation with respect to an event
- Conditional expectation with respect to a random variable

Formal definition

- Conditional expectation with respect to a sub- σ -algebra
- Conditional expectation with respect to a random variable
- Discussion

Conditioning as factorization

Computation

- Basic properties
- See also
- Notes
- References
- External links

Examples

Example 1. Consider the roll of a fair die and let $A = 1$ if the number is even (i.e. 2, 4, or 6) and $A = 0$ otherwise. Furthermore, let $B = 1$ if the number is prime (i.e. 2, 3, or 5) and $B = 0$ otherwise.

	1	2	3	4	5	6
A	0	1	0	1	0	1
B	0	1	1	0	1	0

The unconditional expectation of A is $E[A] = (0 + 1 + 0 + 1 + 0 + 1)/6 = 1/2$. But the expectation of A *conditional* on $B = 1$ (i.e., conditional on the die roll being 2, 3, or 5) is $E[A|B = 1] = (1 + 0 + 0)/3 = 1/3$, and the expectation of A conditional on $B = 0$ (i.e., conditional on the die roll being 1, 4, or 6) is $E[A|B = 0] = (0 + 1 + 1)/3 = 2/3$. Likewise, the expectation of B conditional on $A = 1$ is $E[B|A = 1] = (1 + 0 + 0)/3 = 1/3$, and the expectation of B conditional on $A = 0$ is $E[B|A = 0] = (0 + 1 + 1)/3 = 2/3$.

Example 2. Suppose we have daily rainfall data (mm of rain each day) collected by a weather station on every day of the ten–year (3652–day) period from Jan 1, 1990 to Dec 31, 1999. The unconditional expectation of rainfall for an unspecified day is the average of the rainfall amounts of those 3652 days. The *conditional* expectation of rainfall for an otherwise unspecified day known to be (conditional on being) in the month of March is the average of daily rainfall over all 310 days of the ten–year period that falls in March. And the conditional expectation of rainfall conditional on days dated March 2 is the average of the rainfall amounts that occurred on the ten days with that specific date.

History

The related concept of conditional probability dates back at least to Laplace who calculated conditional distributions. It was Andrey Kolmogorov who in 1933 formalized it using the Radon–Nikodym theorem.^[1] In works of Paul Halmos^[2] and Joseph L. Doob^[3] from 1953, conditional expectation was generalized to its modern definition using sub-σ-algebras.^[4]

Classical definition

Conditional expectation with respect to an event

In classical probability theory the **conditional expectation** of X given an event H (which may be the event $Y = y$ for a random variable Y) is the average of X over all outcomes in H , that is

$$\mathbf{E}(X \mid H) = \frac{\sum_{\omega \in H} X(\omega)}{|H|},$$

Where $|H|$ is the cardinality of H .

The sum above can be grouped by different values of $X(\omega)$, to get a sum over the range \mathcal{X} of X

$$\mathbf{E}(X \mid H) = \sum_{x \in \mathcal{X}} x \frac{|\{\omega \in H \mid X(\omega) = x\}|}{|H|}.$$

In modern probability theory, when H is an event with strictly positive probability, it is possible to give a similar formula. This is notably the case for a discrete random variable Y and for y in the range of Y if the event H is $Y = y$. Let (Ω, \mathcal{F}, P) be a probability space, X is a random variable on that probability space, and $H \in \mathcal{F}$ an event with strictly positive probability $P(H) > 0$. Then the **conditional expectation** of X given the event H is

$$\mathbf{E}(X \mid H) = \frac{\mathbf{E}(1_H X)}{P(H)} = \int_{\mathcal{X}} x \, dP(x \mid H),$$

where \mathcal{X} is the range of X and $P(\cdot \mid H)$ is the probability measure defined, for each set A , as $P(A \mid H) = P(A \cap H)/P(H)$, the conditional probability of A given H .

When $P(H) = 0$ (for instance if Y is a continuous random variable and H is the event $Y = y$, this is in general the case), the Borel–Kolmogorov paradox demonstrates the ambiguity of attempting to define the conditional probability knowing the event H . The above formula shows that this problem transposes to the conditional expectation. So instead one only defines the conditional expectation with respect to a σ -algebra or a random variable.

Conditional expectation with respect to a random variable

If Y is a discrete random variable on the same probability space (Ω, \mathcal{F}, P) having range \mathcal{Y} , then the conditional expectation of X with respect to Y is the random variable $\mathbf{E}(X \mid Y)$ on \mathcal{Y} defined by

$$\mathbf{E}(X \mid Y)(y) = \mathbf{E}(X \mid Y = y).$$

There is a closely related function from Ω to \mathcal{Y} defined by

$$\mathbf{E}(X \mid \sigma(Y))(\omega) = \mathbf{E}(X \mid Y = Y(\omega)).$$

This function, which is different from the previous one, is the conditional expectation of X with respect to the σ -algebra generated by Y . The two are related by

$$\mathbf{E}(X \mid \sigma(Y)) = \mathbf{E}(X \mid Y) \circ Y.$$

As mentioned above, if Y is a continuous random variable, it is not possible to define $\mathbf{E}(X \mid Y)$ by this method. As explained in the [Borel–Kolmogorov paradox](#), we have to specify what limiting procedure produces the set $Y = y$. If the event space Ω has a distance function, then one procedure for doing so is as follows. Define the set $H_y^\epsilon = \{\omega \mid \|Y(\omega) - y\| < \epsilon\}$. Assume that each H_y^ϵ is P -measurable and that $P(H_y^\epsilon) > 0$ for all $\epsilon > 0$. Then conditional expectation with respect to H_y^ϵ is well-defined. Take the limit as ϵ tends to 0 and define

$$g(y) = \lim_{\epsilon \rightarrow 0} \mathbf{E}(X \mid H_y^\epsilon).$$

Replacing this limiting process by the [Radon–Nikodym derivative](#) yields an analogous definition that works more generally.

Formal definition

Conditional expectation with respect to a sub- σ -algebra

Consider the following:

- (Ω, \mathcal{F}, P) is a [probability space](#).
- $X : \Omega \rightarrow \mathbb{R}^n$ is a [random variable](#) on that probability space with finite expectation.
- $\mathcal{H} \subseteq \mathcal{F}$ is a sub- [\$\sigma\$ -algebra](#) of \mathcal{F} .

Since \mathcal{H} is a sub σ -algebra of \mathcal{F} , the function $X : \Omega \rightarrow \mathbb{R}^n$ is usually not \mathcal{H} -measurable, thus the existence of the integrals of the form $\int_H X dP|_{\mathcal{H}}$, where $H \in \mathcal{H}$ and $P|_{\mathcal{H}}$ is the restriction of P to \mathcal{H} , cannot be stated in general. However, the local averages $\int_H X dP$ can be recovered in $(\Omega, \mathcal{H}, P|_{\mathcal{H}})$ with the help of the conditional expectation. A **conditional expectation** of X given \mathcal{H} , denoted as $\mathbf{E}(X \mid \mathcal{H})$, is any \mathcal{H} -[measurable function](#) $\Omega \rightarrow \mathbb{R}^n$ which satisfies:

$$\int_H \mathbf{E}(X \mid \mathcal{H}) dP = \int_H X dP$$

for each $H \in \mathcal{H}$.^[5]

The existence of $\mathbf{E}(X | \mathcal{H})$ can be established by noting that $\mu^X : \mathcal{F} \mapsto \int_{\mathcal{F}} X dP$ for $\mathcal{F} \in \mathcal{F}$ is a finite measure on (Ω, \mathcal{F}) that is absolutely continuous with respect to P . If h is the natural injection from \mathcal{H} to \mathcal{F} , then $\mu^X \circ h = \mu^X|_{\mathcal{H}}$ is the restriction of μ^X to \mathcal{H} and $P \circ h = P|_{\mathcal{H}}$ is the restriction of P to \mathcal{H} . Furthermore, $\mu^X \circ h$ is absolutely continuous with respect to $P \circ h$, because the condition

$$P \circ h(H) = 0 \iff P(h(H)) = 0$$

implies

$$\mu^X(h(H)) = 0 \iff \mu^X \circ h(H) = 0.$$

Thus, we have

$$\mathbf{E}(X | \mathcal{H}) = \frac{d\mu^X|_{\mathcal{H}}}{dP|_{\mathcal{H}}} = \frac{d(\mu^X \circ h)}{d(P \circ h)},$$

where the derivatives are Radon–Nikodym derivatives of measures.

Conditional expectation with respect to a random variable

Consider, in addition to the above,

- A measurable space (U, Σ) , and
- A random variable $Y : \Omega \rightarrow U$.

Let $g : U \rightarrow \mathbb{R}^n$ be a Σ -measurable function such that, for every Σ -measurable function $f : U \rightarrow \mathbb{R}^n$,

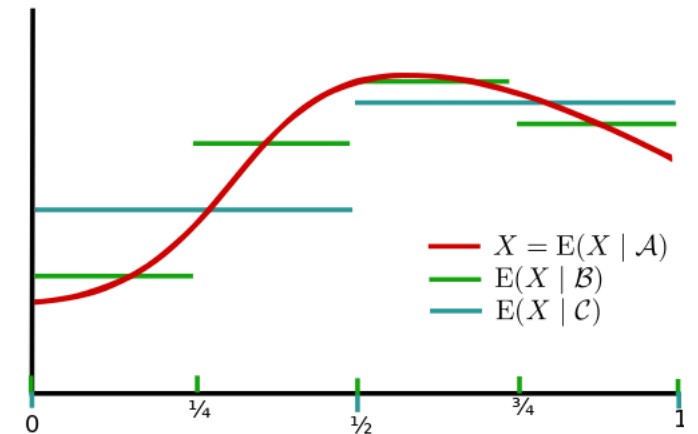
$$\int g(Y)f(Y) dP = \int Xf(Y) dP.$$

Then the random variable $g(Y)$, denoted as $\mathbf{E}(X | Y)$, is a **conditional expectation** of X given Y .

This definition is equivalent to defining the conditional expectation with respect to the sub- σ -field of \mathcal{F} (see above) defined by the pre-image of Σ by Y . If we define

$$\mathcal{H} = Y^{-1}(\Sigma) = \{Y^{-1}(B) : B \in \Sigma\},$$

then



Conditional expectation with respect to a σ -algebra: in this example the probability space (Ω, \mathcal{F}, P) is the $[0, 1]$ interval with the Lebesgue measure. We define the following σ -algebras: $\mathcal{A} = \mathcal{F}$; \mathcal{B} is the σ -algebra generated by the intervals with end-points $0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1$; and \mathcal{C} is the σ -algebra generated by the intervals with end-points $0, \frac{1}{2}, 1$. Here the conditional expectation is effectively the average over the minimal sets of the σ -algebra.

$$\mathbf{E}(X \mid Y) = \mathbf{E}(X \mid \mathcal{H}) = \frac{d(\mu^X \circ Y^{-1})}{d(P \circ Y^{-1})} \circ Y.$$

Discussion

- This is not a constructive definition; we are merely given the required property that a conditional expectation must satisfy.
 - The definition of $\mathbf{E}(X \mid \mathcal{H})$ may resemble that of $\mathbf{E}(X \mid H)$ for an event H but these are very different objects. The former is a \mathcal{H} -measurable function $\Omega \rightarrow \mathbb{R}^n$, while the latter is an element of \mathbb{R}^n . Evaluating the former at H yields the latter.
 - Existence of a conditional expectation function may be proven by the Radon–Nikodym theorem. A sufficient condition is that the (unconditional) expected value for X exists.
 - Uniqueness can be shown to be almost sure: that is, versions of the same conditional expectation will only differ on a set of probability zero.
- The σ -algebra \mathcal{H} controls the "granularity" of the conditioning. A conditional expectation $\mathbf{E}(X \mid \mathcal{H})$ over a finer (larger) σ -algebra \mathcal{H} retains information about the probabilities of a larger class of events. A conditional expectation over a coarser (smaller) σ -algebra averages over more events.

Conditioning as factorization

In the definition of conditional expectation that we provided above, the fact that Y is a *real* random element is irrelevant. Let (U, Σ) be a measurable space, where Σ is a σ -algebra on U . A *U -valued random element* is a measurable function $Y: \Omega \rightarrow U$, i.e. $Y^{-1}(B) \in \mathcal{F}$ for all $B \in \Sigma$. The *distribution* of Y is the probability measure $P_Y: \Sigma \rightarrow \mathbb{R}$ defined as the pushforward measure Y_*P , that is, such that $P_Y(B) = P(Y^{-1}(B))$.

Theorem. If $X: \Omega \rightarrow \mathbb{R}$ is an integrable random variable, then there exists a unique integrable random element $\mathbf{E}(X \mid Y): U \rightarrow \mathbb{R}$, defined P_Y almost surely, such that

$$\int_{Y^{-1}(B)} X \, dP = \int_B \mathbf{E}(X \mid Y) \, dP_Y,$$

for all $B \in \Sigma$.

Proof sketch. Let $\mu: \Sigma \rightarrow \mathbb{R}$ be such that $\mu(B) = \int_{Y^{-1}(B)} X \, dP$. Then μ is a signed measure which is absolutely continuous with respect to P_Y . Indeed $P_Y(B) = 0$ means exactly that $P(Y^{-1}(B)) = 0$, and since the integral of an integrable function on a set of probability 0 is 0, this proves absolute continuity. The Radon–Nikodym theorem then proves the existence of a density of μ with respect to P_Y . This density is $\mathbf{E}(X \mid Y)$. \square

Comparing with conditional expectation with respect to sub- σ -algebras, it holds that

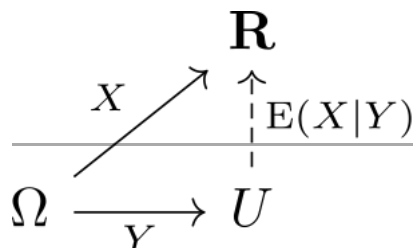
$$\mathbf{E}(X \mid Y) \circ Y = \mathbf{E}(X \mid Y^{-1}(\Sigma)).$$

We can further interpret this equality by considering the abstract change of variables formula to transport the integral on the right hand side to an integral over Ω :

$$\int_{Y^{-1}(B)} X \, dP = \int_{Y^{-1}(B)} (\mathbf{E}(X \mid Y) \circ Y) \, dP.$$

The equation means that the integrals of X and the composition $\mathbf{E}(X \mid Y) \circ Y$ over sets of the form $Y^{-1}(B)$, for $B \in \Sigma$, are identical.

This equation can be interpreted to say that the following diagram is commutative on average.



Computation

When X and Y are both discrete random variables, then the conditional expectation of X given the event $Y = y$ can be considered as function of y for y in the range of Y :

$$\mathbf{E}(X \mid Y = y) = \sum_{x \in \mathcal{X}} x P(X = x \mid Y = y) = \sum_{x \in \mathcal{X}} x \frac{P(X = x, Y = y)}{P(Y = y)},$$

where \mathcal{X} is the range of X .

If X is a continuous random variable, while Y remains a discrete variable, the conditional expectation is

$$\mathbf{E}(X \mid Y = y) = \int_{\mathcal{X}} x f_X(x \mid Y = y) \, dx,$$

with $f_X(x \mid Y = y) = \frac{f_{X,Y}(x, y)}{P(Y = y)}$ (where $f_{X,Y}(x, y)$ gives the joint density of X and Y) being the conditional density of X given $Y = y$.

If both X and Y are continuous random variables, then the conditional expectation is

$$\mathbf{E}(X \mid Y = y) = \int_{\mathcal{X}} x f_{X|Y}(x \mid y) \, dx,$$

where $f_{X|Y}(x \mid y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}$ (where $f_Y(y)$ gives the density of Y).

Basic properties

All the following formulas are to be understood in an almost sure sense. The σ -algebra \mathcal{H} could be replaced by a random variable Z .

- Pulling out independent factors:
 - If X is independent of \mathcal{H} , then $\mathbf{E}(X \mid \mathcal{H}) = \mathbf{E}(X)$.

Proof

Let $B \in \mathcal{H}$. Then X is independent of $\mathbf{1}_B$, so we get that

$$\int_B X \, dP = \mathbf{E}(X \mathbf{1}_B) = \mathbf{E}(X) \mathbf{E}(\mathbf{1}_B) = \mathbf{E}(X) P(B) = \int_B \mathbf{E}(X) \, dP.$$

Thus the definition of conditional expectation is satisfied by the constant random variable $\mathbf{E}(X)$, as desired.

- ▪ If X is independent of $\sigma(Y, \mathcal{H})$, then $\mathbf{E}(XY \mid \mathcal{H}) = \mathbf{E}(X) \mathbf{E}(Y \mid \mathcal{H})$. Note that this is not necessarily the case if X is only independent of \mathcal{H} and of Y .
- If X, Y are independent, \mathcal{G}, \mathcal{H} are independent, X is independent of \mathcal{H} and Y is independent of \mathcal{G} , then $\mathbf{E}(\mathbf{E}(XY \mid \mathcal{G}) \mid \mathcal{H}) = \mathbf{E}(X) \mathbf{E}(Y) = \mathbf{E}(\mathbf{E}(XY \mid \mathcal{H}) \mid \mathcal{G})$.
- Stability:
 - If X is \mathcal{H} -measurable, then $\mathbf{E}(X \mid \mathcal{H}) = X$.
 - If Z is a random variable, then $\mathbf{E}(f(Z) \mid Z) = f(Z)$. In its simplest form, this says $\mathbf{E}(Z \mid Z) = Z$.
- Pulling out known factors:
 - If X is \mathcal{H} -measurable, then $\mathbf{E}(XY \mid \mathcal{H}) = X \mathbf{E}(Y \mid \mathcal{H})$.
 - If Z is a random variable, then $\mathbf{E}(f(Z)Y \mid Z) = f(Z) \mathbf{E}(Y \mid Z)$.
- Law of total expectation: $\mathbf{E}(\mathbf{E}(X \mid \mathcal{H})) = \mathbf{E}(X)$.
- Tower property:
 - For sub- σ -algebras $\mathcal{H}_1 \subset \mathcal{H}_2 \subset \mathcal{F}$ we have $\mathbf{E}(\mathbf{E}(X \mid \mathcal{H}_2) \mid \mathcal{H}_1) = \mathbf{E}(X \mid \mathcal{H}_1)$.

- A special case is when Z is a \mathcal{H} -measurable random variable. Then $\sigma(Z) \subset \mathcal{H}$ and thus $E(E(X | \mathcal{H}) | Z) = E(X | Z)$.
- Doob martingale property: the above with $Z = E(X | \mathcal{H})$ (which is \mathcal{H} -measurable), and using also $E(Z | Z) = Z$, gives $E(X | E(X | \mathcal{H})) = E(X | \mathcal{H})$.
- For random variables X, Y we have $E(E(X | Y) | f(Y)) = E(X | f(Y))$.
- For random variables X, Y, Z we have $E(E(X | Y, Z) | Y) = E(X | Y)$.
- Linearity: we have $E(X_1 + X_2 | \mathcal{H}) = E(X_1 | \mathcal{H}) + E(X_2 | \mathcal{H})$ and $E(aX | \mathcal{H}) = aE(X | \mathcal{H})$ for $a \in \mathbb{R}$.
- Positivity: If $X \geq 0$ then $E(X | \mathcal{H}) \geq 0$.
- Monotonicity: If $X_1 \leq X_2$ then $E(X_1 | \mathcal{H}) \leq E(X_2 | \mathcal{H})$.
- Monotone convergence: If $0 \leq X_n \uparrow X$ then $E(X_n | \mathcal{H}) \uparrow E(X | \mathcal{H})$.
- Dominated convergence: If $X_n \rightarrow X$ and $|X_n| \leq Y$ with $Y \in L^1$, then $E(X_n | \mathcal{H}) \rightarrow E(X | \mathcal{H})$.
- Fatou's lemma: If $E(\inf_n X_n | \mathcal{H}) > -\infty$ then $E(\liminf_{n \rightarrow \infty} X_n | \mathcal{H}) \leq \liminf_{n \rightarrow \infty} E(X_n | \mathcal{H})$.
- Jensen's inequality: If $f: \mathbb{R} \rightarrow \mathbb{R}$ is a convex function, then $f(E(X | \mathcal{H})) \leq E(f(X) | \mathcal{H})$.
- Conditional variance: Using the conditional expectation we can define, by analogy with the definition of the variance as the mean square deviation from the average, the conditional variance
 - Definition: $\text{Var}(X | \mathcal{H}) = E((X - E(X | \mathcal{H}))^2 | \mathcal{H})$
 - Algebraic formula for the variance: $\text{Var}(X | \mathcal{H}) = E(X^2 | \mathcal{H}) - (E(X | \mathcal{H}))^2$
 - Law of total variance: $\text{Var}(X) = E(\text{Var}(X | \mathcal{H})) + \text{Var}(E(X | \mathcal{H}))$.
- Martingale convergence: For a random variable X , that has finite expectation, we have $E(X | \mathcal{H}_n) \rightarrow E(X | \mathcal{H})$, if either $\mathcal{H}_1 \subset \mathcal{H}_2 \subset \dots$ is an increasing series of sub- σ -algebras and $\mathcal{H} = \sigma(\bigcup_{n=1}^{\infty} \mathcal{H}_n)$ or if $\mathcal{H}_1 \supset \mathcal{H}_2 \supset \dots$ is a decreasing series of sub- σ -algebras and $\mathcal{H} = \bigcap_{n=1}^{\infty} \mathcal{H}_n$.
- Conditional expectation as L^2 -projection: If X, Y are in the Hilbert space of square-integrable real random variables (real random variables with finite second moment) then
 - for \mathcal{H} -measurable Y , we have $E(Y(X - E(X | \mathcal{H}))) = 0$, i.e. the conditional expectation $E(X | \mathcal{H})$ is in the sense of the $L^2(P)$ scalar product the orthogonal projection from X to the linear subspace of \mathcal{H} -measurable functions. (This allows to define and prove the existence of the conditional expectation based on the Hilbert projection theorem.)
 - the mapping $X \mapsto E(X | \mathcal{H})$ is self-adjoint: $E(X E(Y | \mathcal{H})) = E(E(X | \mathcal{H}) E(Y | \mathcal{H})) = E(E(X | \mathcal{H}) Y)$
- Conditioning is a contractive projection of L^p spaces $L^p(\Omega, \mathcal{F}, P) \rightarrow L^p(\Omega, \mathcal{H}, P)$. i.e., $E(|E(X | \mathcal{H})|^p) \leq E(|X|^p)$ for any $p \geq 1$.
- Doob's conditional independence property:^[6] If X, Y are conditionally independent given Z , then $P(X \in B | Y, Z) = P(X \in B | Z)$ (equivalently, $E(1_{\{X \in B\}} | Y, Z) = E(1_{\{X \in B\}} | Z)$).

See also

- | | | |
|--------------------------------------------------|-----------------------------------|---------------------------------------------------------------|
| ▪ <u>Non-commutative conditional expectation</u> | ▪ <u>Law of total expectation</u> | ▪ <u>Law of total cumulance</u> (generalizes the other three) |
| ▪ <u>Law of total probability</u> | ▪ <u>Law of total variance</u> | |

- [Conditioning \(probability\)](#)
- [Joint probability distribution](#)
- [Disintegration theorem](#)
- [Doob–Dynkin lemma](#)
- [Factorization lemma](#)

Notes

1. [Kolmogorov, Andrey](#) (1933). *Grundbegriffe der Wahrscheinlichkeitsrechnung* (in German). Berlin: Julius Springer. p. 46.
 - Translation: [Kolmogorov, Andrey](#) (1956). *Foundations of the Theory of Probability* (<http://www.mathematik.com/Kolmogorov/index.html>) (2nd ed.). New York: Chelsea. ISBN 0-8284-0023-7.
2. Oxtoby, J. C. (1953). "Review: *Measure theory*, by P. R. Halmos" (<http://www.ams.org/journals/bull/1953-59-01/S0002-9904-1953-09662-8/S0002-9904-1953-09662-8.pdf>) (PDF). *Bull. Amer. Math. Soc.* **59** (1): 89–91. doi:10.1090/s0002-9904-1953-09662-8 (<https://doi.org/10.1090%2Fs0002-9904-1953-09662-8>).
3. J. L. Doob (1953). *Stochastic Processes*. John Wiley & Sons. ISBN 0-471-52369-0.
4. Olav Kallenberg: *Foundations of Modern Probability*. 2. edition. Springer, New York 2002, ISBN 0-387-95313-2, S. 573.
5. [Billingsley, Patrick](#) (1995). "Section 34. Conditional Expectation". *Probability and Measure* (3rd ed.). John Wiley & Sons. p. 445. ISBN 0-471-00710-2.
6. Kallenberg, Olav (2001). *Foundations of Modern Probability* (2nd ed.). York, PA, USA: Springer. p. 110. ISBN 0-387-95313-2.

References

- [William Feller](#), *An Introduction to Probability Theory and its Applications*, vol 1, 1950, page 223
- Paul A. Meyer, *Probability and Potentials*, Blaisdell Publishing Co., 1966, page 28
- [Grimmett, Geoffrey](#); [Stirzaker, David](#) (2001). *Probability and Random Processes* (3rd ed.). Oxford University Press. ISBN 0-19-857222-0., pages 67–69

External links

- [Ushakov, N.G.](#) (2001) [1994], "Conditional mathematical expectation" (<https://www.encyclopediaofmath.org/index.php?title=c/c024500>), in [Hazewinkel, Michiel](#) (ed.), *Encyclopedia of Mathematics*, Springer Science+Business Media B.V. / Kluwer Academic Publishers, ISBN 978-1-55608-010-4

Retrieved from "https://en.wikipedia.org/w/index.php?title=Conditional_expectation&oldid=903905131"

This page was last edited on 28 June 2019, at 17:41 (UTC).

Text is available under the Creative Commons Attribution-ShareAlike License; additional terms may apply. By using this site, you agree to the [Terms of Use](#) and [Privacy Policy](#). Wikipedia® is a registered trademark of the [Wikimedia Foundation, Inc.](#), a non-profit organization.