Chapter 5. Joint distribution

5.1 Jointly distributed r.v.'s

- (Joint) cumulative distribution function (cdf) of *X* and *Y* is defined as the bivariate function

$$F(a,b) = P\{X \le a, Y \le b\}, -\infty < a, b < \infty$$

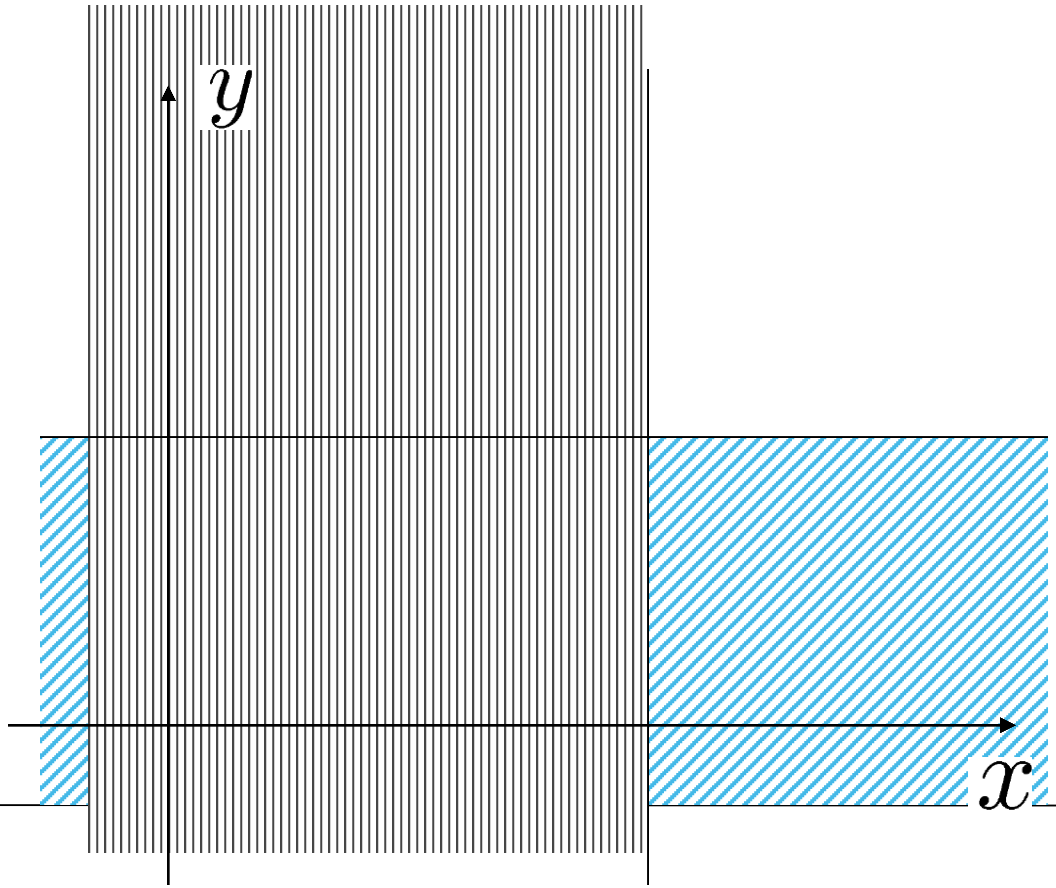
Marginal distributions of X and Y are given by

$$\begin{aligned} F_X(a) &= P\{X \le a\} = P\{X \le a, Y \le \infty\} \\ &= \lim_{b \to \infty} P\{X \le a, Y \le b\} = F(a, \infty) \end{aligned}$$

$$\text{Similarly, } F_Y(b) = F(\infty, b)$$

$$P\{X \le a, \text{ or } Y \le b\} = F_X(a) + F_Y(b) - F(a,b)$$

$$P\{a_1 < X \le a_2, b_1 < Y \le b_2\} = F(a_2,b_2) + F(a_1,b_1) - F(a_1,b_2) - F(a_2,b_1)$$



Joint probability mass functions: Let *X* and *Y* be discrete random variables taking on values $x_1, x_2, \ldots,$ and $y_1, y_2, \ldots,$ respectively. The joint probability mass function of (*X*,*Y*) is $p(x_i, y_j) = P\{X = x_i, Y = y_j\}$

The marginal pmf can be computed as

$$p_X(x_i) = \sum_j p(x_i, y_j), p_Y(y_j) = \sum_i p(x_i, y_j)$$

Joint probability density function: random variables *X* and *Y* are said to be jointly continuous if there exists a function f(x,y) defined for all real x and y, having the property that for every set $C \subset R^2$ (*C* is a set in the two-dimensional plane)

$$P\{(X,Y) \in C\} = \int\int_C f(x,y)\,dxdy$$

f(x,y) is called joint density function of *X* and *Y*.

The marginal pdf for *X* is defined as: $f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy$

Example : A fair coin is tossed three times independently; let *X* denote the number of heads on the first toss and *Y* denote the total number of heads. Find the joint probability mass function of *X* and *Y*, together with the marginal pmf of *X* and *Y*.

Solution: The joint and marginal pmf is given in the following table:

	Y				
X	0	1	2	3	p(x)
0	1/8	1/4	1/8	0	1/2
1	0	1/8	1/4	1/8	1/2
p(y)	1/8	3/8	3/8	1/8	

One of the most important discrete joint distributions is *Multinomial distribution*: A sequence of *n* independent and identical experiments is performed, each resulting in any one of *r* possible outcomes, with respective probabilities $p_1, p_2, \ldots, p_r, \; \sum_{i=1}^{r} p_i = 1.$

Let X_i denote the number of the *n* experiments that result in outcome *i*, then

$$P\{X_1 = n_1, X_2 = n_2, \ldots, X_r = n_r\} = \frac{n!}{n_1! n_2! \ldots n_r!} p_1^{n_1} p_2^{n_2} \ldots p_r^{n_r}$$

Example : Suppose that a fair die is rolled 9 times. What is the probability that 1 appears three times, 2 and 3 twice each, 4 and 5 once each and 6 not at all?

Solution: $$\frac{9!}{3!2!2!1!1!0!}(\frac{1}{6})^9$$

Example : The joint density function of X and Y is given by
$$f(x,y) = \begin{cases} 2e^{-x}e^{-2y} & 0 < x < \infty, 0 < y < \infty \\ 0 & \text{otherwise} \end{cases}$$

Compute (a) $P\{X>1,Y<1\}$ (b) $P\{X<a\}$ (c) $P\{X<Y\}$

Solution:

(a)

Example : The joint density function of *X* and *Y* is given by

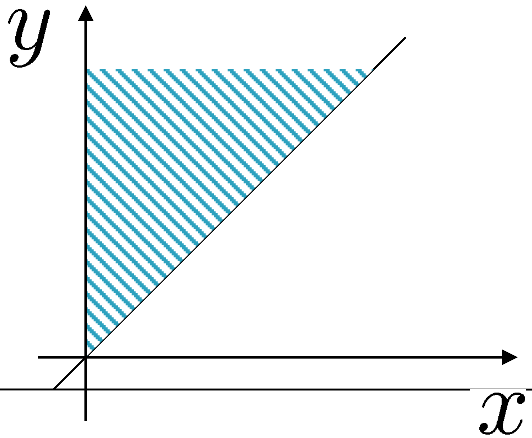
$$f(x,y) = \begin{cases} 2e^{-x}e^{-2y} & 0 < x < \infty, 0 < y < \infty \\ 0 & \text{otherwise} \end{cases}$$

Compute (a) $P\{X>1, Y<1\}$ (b) $P\{X<a\}$ (c) $P\{X<Y\}$

Solution:

(b)

(c)



Independent Random Variables

Definition: X and Y are said to be independent if for any two sets of real number A and B,

$$P\{X \in A, Y \in B\} = P\{X \in A\}P\{Y \in B\} \qquad (*)$$

It can be shown that (*) will follow if and only if for all *a* and *b* $F(a,b) = F_X(a)F_Y(b)$

When *X* and *Y* are discrete, it is equivalent to (this you can prove) $p(x,y) = p_X(x)p_Y(y) \quad \text{for all } x, y$

In continuous case, it is equivalent to (this you can prove too) $f(x,y) = f_X(x)f_Y(y) \quad \text{for all } x, y$

Example: A fair die is rolled twice. Let X be the outcome of the first roll, and Z be the sum of the two rolls. Are X and Z independent?

Solution: We showed in a previous example (Chapter 2) that the events *{X=4}* and *{Z=6}* are dependent, while *{X=4}* and *{Z=7}* are independent. Thus.....

Example: (try to memorize the result) Suppose that the number of people that enter a post office on a given day is a Poisson random variable with parameter λ . Show that if each person that enters the post office is a male with probability p and a female with probability 1-p, then the number of males and females entering the post office are independent Poisson random variables with respective parameters λp and $\lambda(1-p)$

Proposition: The continuous (discrete) random variables X and Y are independent if and only if their joint probability density (mass) function can be expressed as

$$f_{XY}(xy) = h(x)g(y)$$

for all real numbers x,y.

Example. If the joint density function of X and Y is

(a) $f(x,y) = 6e^{-2x}e^{-3y} \quad 0 < x < \infty, 0 < y < \infty$

(b) $f(x,y) = 24xy \quad 0 < x < 1, 0 < y < 1, 0 < x + y < 1$

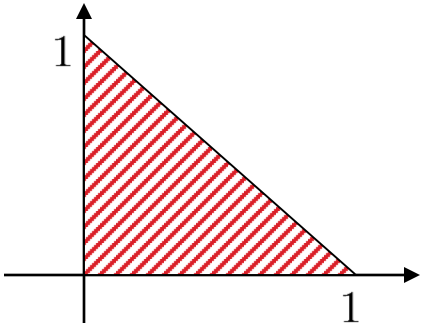
and equal to 0 otherwise, are X and Y independent?

Solution:

(a) Independent

(b) Dependent

positive domain in (b)



Example: Suppose X_1, X_2, X_3 are independent and distributed as $exp(\lambda_1), exp(\lambda_2), exp(\lambda_3)$ respectively. Find the probability that $\min(X_1, X_2, X_3) = X_2$

General Definition of Independence: $X_1, X_2, \ldots, X_n$ are said to be independent if, for all sets $A_1, A_2, \ldots, A_n$,

$$P\{X_1 \in A_1, X_2 \in A_2, \ldots, X_n \in A_n\} = \prod_{i=1}^{n} P\{X_i \in A_i\}$$

Sums of Independent Random Variables

It is often important to calculate the distribution of *X+Y* given the distribution of *X* and *Y* when *X* and *Y* are independent.

$$\begin{aligned}
F_{X+Y}(a) &= P\{X+Y \leq a\} \\
&= \iint_{x+y\leq a} f_X(x) f_Y(y)\,dx\,dy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{a-y} f_X(x) f_Y(y)\,dx\,dy \\
&= \int_{-\infty}^{\infty} f_Y(y) \int_{-\infty}^{a-y} f_X(x)\,dx\,dy \\
&= \int_{-\infty}^{\infty} F_X(a-y) f_Y(y)\,dy
\end{aligned}$$

Differentiating on both sides, we get

$$\begin{aligned} f_{X+Y}(a) &= \frac{d}{da}\int_{-\infty}^{\infty} F_X(a-y) f_Y(y) dy \\ &= \int_{-\infty}^{\infty} \frac{d}{da} F_X(a-y) f_Y(y) dy \\ &= \int_{-\infty}^{\infty} f_X(a-y) f_Y(y) dy \end{aligned}$$

The last integral above is usually called convolution of f_X and f_Y

Proposition: If $X_1, X_2, \ldots, X_n$ are independent random variables that are normally distributed with parameters $\mu_i, \sigma_i^2, i = 1, 2, \ldots, n,$ then $\sum_{i=1}^n X_i$ is normally distributed with parameters $\sum_{i=1}^n \mu_i$ and $\sum_{i=1}^n \sigma_i^2$

(Other important results) If X and Y are independent Poisson random variables with respective parameters $\lambda_1 \text{ and } \lambda_2$, then X+Y is a Poisson random variable with parameters $\lambda_1 + \lambda_2$

If X and Y are independent binomial random variables with respective parameters $(n, p) \text{ and } (m, p)$, then X+Y is a binomial random variable with parameters $(n + m, p)$

5.2 Conditional Distributions

Discrete Case

Definition: If X and Y are jointly distributed discrete random variables, the conditional probability that $X = x$ given that $Y = y$ *(conditional pmf)* is

$$p_{X|Y}(x|y) = P\{X = x|Y = y\} = \frac{p(x,y)}{p_Y(y)}$$

Remark: If X and Y are independent random variables, then the conditional probability mass function is the same as the unconditional one.

Example : Suppose that *p(x,y),* the joint probability mass function of X and Y, is given by p(0,0)=.4 p(0,1)=.2 p(1,0)=.1 p(1,1)=.3 Calculate the conditional probability mass function of X, given that Y=1

Solution:

Example: If X and Y are independent Poisson random variables with parameters $\lambda_1 \text{ and } \lambda_2$ respectively, calculate the conditional distribution of X, given that $X+Y=n$.

Solution:

Continuous Case:

Definition: If *X* and *Y* have a joint probability density function *f*(*x*,*y*), then the conditional pdf of *X*, given that *Y*=*y* is defined for all values of *y* such that $f_Y(y) > 0$ by $$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}$$

If X and Y are independent, $f_{X|Y}(x|y) = f_X(x)$

If *X* and *Y* are jointly continuous, then for any set *A*,

$$P\{X \in A|Y = y\} = \int_A f_{X|Y}(x|y)dx$$

In particular, let $A = (-\infty, a]$, we can define the conditional cdf of *X* given that *Y*=*y* by

$$F_{X|Y}(a|y) = P\{X \leq a|Y = y\}$$

Example : The joint density function of X and Y is given by

$$f(x,y) = \begin{cases} \frac{15}{2}x(2-x-y) & 0 < x, y < 1 \\ 0 & \text{otherwise} \end{cases}$$

Find f(x|y)

Solution:

Example : The joint density function of X and Y is given.

Compute *P{X>1|Y=y}*

$$f(x,y) = \begin{cases} \frac{e^{-x/y}e^{-y}}{y} & 0 < x, y < \infty \\ 0 & \text{otherwise} \end{cases}$$

5.3 Joint Distribution of Functions of Random Variables

Let $X_1 \text{ and } X_2$ be jointly continuous random variables with joint pdf f_{XY}. We want to compute the density function of $Y_1 = g_1(X_1, X_2) \text{ and } Y_2 = g_2(X_1, X_2)$

Assume that the functions $g_1 \text{ and } g_2$ satisfy the following conditions:

1. $y_1 = g_1(x_1, x_2) \text{ and } y_2 = g_2(x_1, x_2)$ can be uniquely solved for $x_1 \text{ and } x_2$ to get, say, $x_1 = h_1(y_1, y_2), x_2 = h_2(y_1, y_2)$
2. The functions $g_1 \text{ and } g_2$ have continuous partial derivatives and the determinant of the following determinant $J(x_1, x_2) = \begin{vmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} \end{vmatrix} \neq 0$

Then $f_{Y_1,Y_2}(y_1, y_2) = f_{X_1,X_2}(x_1, x_2)|J(x_1, x_2)|^{-1} \text{ where}$

$$x_1 = h_1(y_1, y_2), x_2 = h_2(y_1, y_2)$$

5.4 Expectation, covariance, conditional expectation

Discrete $E[X] = \sum_{x_i} x_i P\{X = x_i\}$

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

Since E[X] is a weighted average of possible values of X,

$$\text{if } P\{a \le X \le b\} = 1, \text{ then } a \le E[X] \le b.$$

Recall

$$E[g(X)] = \sum_{x_i} g(x_i) p(x_i), E[g(X)] = \int_{-\infty}^{\infty} g(x) f(x)\, dx$$

A two-dimensional analog is the following:

Proposition.

$$E[g(X,Y)] = \sum_x \sum_y g(x,y)p(x,y)$$

$$E[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f(x,y)\,dxdy$$

An important implication is

$$\begin{aligned}
E[X+Y] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)f(x,y)dxdy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf(x,y)dxdy + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yf(x,y)dxdy \\
&= \int_{-\infty}^{\infty} xf_X(x)dx + \int_{-\infty}^{\infty} yf_Y(y)dy \\
&= E[X] + E[Y]
\end{aligned}$$

By induction, we have

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

Example: (Sample mean). Let $X_1, \ldots, X_n$ be independent and identically distributed (i.i.d.) random variables having cdf *F* and expected value μ. Such a sequence of r.v.'s is said to be a *random sample* from *F*. The sample mean $\bar{X}$ is defined by $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$

Compute its expectation.

Solution:

Example: (Mean of a negative binomial random variable.) If independent trials, each having a constant probability p of being a success are performed, determine the expected number of trials required to amass a total of r successes.

Solution: The negative binomial r.v. with parameter r and p is a sum of r i.i.d. geometric(p) r.v.'s. So the expectation of the negative binomial r.v. is the sum of the expectation of r geometric r.v.'s. So the answer is r/p.

Example: Consider the process of randomly tossing identical balls into b bins, numbered 1, 2,..., b. The tosses are independent, and on each toss the ball is equally likely to end up in any bin. The probability that a tossed ball lands in any given bin is $1/b$. Thus, the ball-tossing process is a sequence of Bernoulli trials with a probability $1/b$ of success, where success means that the ball falls in the given bin.

(a) *How many balls fall in a given bin (if a total of n balls are tossed)?*

(b) How many balls must one toss, on the average, until a given bin contains a ball? The number of tosses until the given bin receives a ball follows the geometric distribution with probability $1/b$ and thus the expected number of tosses until success is $1/(1/b) = b$.

(c) How many balls must one toss until every bin contains at least one ball?

Covariance

The covariance between two random variables is a measure of how they are related.

Definition: The covariance between X and Y, denoted by $Cov(X,Y)$, is defined by $Cov(X,Y)=E[(X-E[X])(Y-E[Y])]$.

Interpretation: When $Cov(X,Y)>0$, higher than expected values of X tend to occur together with higher than expected values of Y. When $Cov(X,Y)<0$, higher than expected values of X tend to occur together with lower than expected values of Y.

By expanding the right hand side of the definition of the covariance, we see that

$$
\begin{aligned}
& Cov(X,Y) \\
= \quad & E[(X-E[X])(Y-E[Y])] = E\{XY - E[X]Y - XE[Y] + E[X]E[Y]\} \\
= \quad & E[XY] - E[X]E[Y] - E[X]E[Y] + E[X]E[Y] = E[XY] - E[X]E[Y]
\end{aligned}
$$

If X and Y are independent, then

$$
\begin{aligned}
E[XY] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)dxdy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_X(x)f_Y(y)dxdy \\
&= E[X]E[Y]
\end{aligned}
$$

Definition: If $Cov(X,Y)=0$, we say X and Y are uncorrelated. If $Cov(X,Y)>0$, we say X and Y are positively correlated. If $Cov(X,Y)<0$, we say X and Y are negatively correlated.

So the previous calculation tells us that independence implies uncorrelatedness.

We have

Proposition: If X and Y are independent, then for any functions of g and h, g(X) and h(Y) are independent.

Proposition

$$(i) \quad Cov(X,Y) = Cov(Y,X)$$

$$(ii) \quad Cov(X,X) = Var(X)$$

$$(iii) \quad Cov(aX,bY) = abCov(X,Y)$$

$$(iv) \quad Cov(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j) = \sum_{i=1}^{n}\sum_{j=1}^{m} Cov(X_i, Y_j)$$

$$(v) \quad Var(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} Var(X_i) + 2\sum_{i<j} Cov(X_i, X_j)$$

(vi) if $X_1, X_2, \ldots, X_n$ are pairwise independent,

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

Correlation: $\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$

The correlation is always between -1 and 1. If *X* and *Y* are independent, then $\rho(X,Y)=0$ .but the converse is not true. Generally, the correlation (as well as covariance) is a measure of the degree of linear dependence between *X* and *Y*.

Note that for *a*>0,*b*>0 ,

$$\rho(aX,bY) = \frac{Cov(aX,bY)}{\sqrt{Var(aX)Var(bY)}} = \frac{abCov(X,Y)}{\sqrt{a^2b^2Var(X)Var(Y)}} = \rho(X,Y)$$

Example : Let $X_1, \ldots, X_n$ be independent and identically distributed random variables having expected value μ and variance σ^2 . and let $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ be the sample mean. The random variable $S^2 = \sum_{i=1}^{n} \frac{(X_i - \bar{X})^2}{n-1}$ is called the sample variance. Find (a) $E(\bar{X})$ and $Var(\bar{X})$ and (b) $E[S^2]$

Example. Compute the variance of a binomial random variable X with parameters n and p.

Solution:

Example . Let I_A and I_B be indicator variables for the events *A* and *B*. Find $Cov(I_A, I_B)$

Solution:

Thus two events are independent if and only if the corresponding indicator variables are uncorrelated. In other words, for indicator variables, independence and uncorrelatedness are equivalent.

Example. Let $X_1, \dots, X_n$ be independent and identically distributed random variables having variance σ^2. Show that $Cov(X_i - \bar{X}, \bar{X}) = 0$

Bivariate normal distribution

Definition: The joint density for a bivariate normal distribution is

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} exp\left\{-\frac{1}{2(1-\rho^2)}\left[(\frac{x-\mu_x}{\sigma_x})^2 + (\frac{y-\mu_y}{\sigma_y})^2 - 2\rho\frac{(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y}\right]\right\}$$

Remarks on bivariate normal random variables (X,Y):

(a) Marginally, $X \sim N(\mu_x, \sigma_x^2), Y \sim N(\mu_y, \sigma_y^2)$

(b) Conditionally, $X|Y = y \sim N(\mu_x + \rho\frac{\sigma_x}{\sigma_y}(y - \mu_y), \sigma_x^2(1 - \rho^2))$

(c) $Cov(X, Y) = \rho\sigma_x\sigma_y$

(d) Linear combinations of X and Y are normal random variables, even though X and Y are not independent when $\rho \neq 0$

(e) Two normal random variables are independent iff they are uncorrelated.

Example: For bivariate normal random variables X and Y with parameters $(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$, find *P(X<Y)*

Solution:

Conditional Expectation

Recall that if X and Y are joint discrete random variables,

It is natural to define, the conditional expectation of *X* given *Y* = *y*

$$E[X|Y=y] = \sum_x xP(X=x|Y=y) = \sum_x xp_{X|Y}(x|y)$$

for continuous random variables:

The conditional expectation of *X*, given that *Y* = *y* , is

$$E[X|Y = y] = \int_{-\infty}^{\infty} x f_{X|Y}(x|y) dx$$

Example. Let *X* and *Y* have the joint pdf

$$f(x,y) = \frac{e^{-x/y}e^{-y}}{y}, 0 < x, y < \infty$$

Find the conditional expectation $E(X \mid Y = y)$.

Remark: 1) You don't need to calculate the marginal density first, as I showed in class.

2) p(x|y) is the density for *exp(1/y),* so you can directly get expectation is y.

Conditional Variance (for your information only)

The conditional variance of $X|Y = y$ is the expected squared difference of the random variable X and its conditional mean, conditioning on the event $Y = y$:

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

Similar to the unconditional case, we can show

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