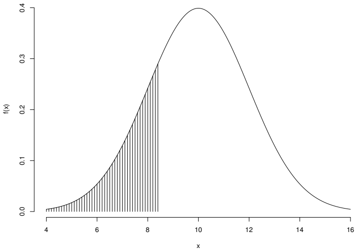
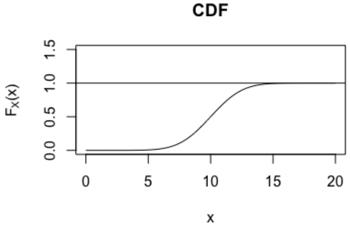
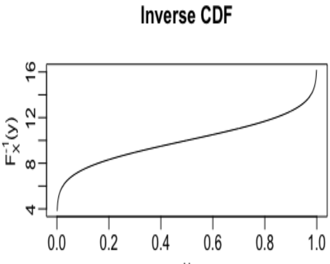
Continuous random variables: review





Density $P(X \leq x) = F_X(x) = \int_{-\infty}^{x} f_X(x)dx$



Same rules for continuous and discrete random variables I

$$f_X(x) \to \begin{cases} f_X(x) \geq 0, \ \int f_X(x)dx = 1 & \text{continuous} \\ f_X(x) \geq 0, \text{for } x \in R_X, \ \sum_{x \in R_X} f_X(x) = 1 & \text{discrete} \end{cases}$$

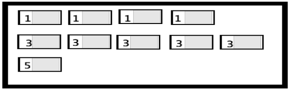
$$E(X) = \begin{cases} \int x f_X(x)dx & \text{continuous} \\ \sum_{x \in R_x} x f_X(x) & \text{discrete} \end{cases}$$

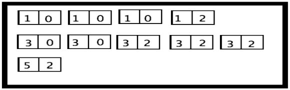
$$Var(x) = \begin{cases} \int [x - E(X)]^2 f_X(x)dx & \text{continuous} \\ \sum_{x \in R_x} [x - E(X)]^2 f_X(x) & \text{discrete} \end{cases}$$

$$Eg(X) = \begin{cases} \int g(x) f_X(x)dx & \text{continuous} \\ \sum_{x \in R_x} g(x) f_X(x) & \text{discrete} \end{cases}$$

Multiple random variables and joint distributions I

Back to box of cards. Original box (4 1's, 5 3's, 1 5), with additional number on right side of each card:





Draw card at random. Define two random variables:
$X-$ value on left. $Y-$ value on right. $R_x = \{1, 3, 5\}, R_y = \{0, 2\}$.
We can fill a two way table with the *joint distribution* of X and Y:
$f_{X,Y}(x, y) = P(X = x \text{ and } Y = y), x \in R_x, y \in R_y$

Y \X	1	3	5	Marg Y
0	3/10	2/10	0	5/10
2	1/10	3/10	1/10	5/10
Marg X	4/10	5/10	1/10	1

Multiple random variables and joint distributions II

Marginal distribution of X sum over all possible values of Y:

$f_X(x) = \sum_{y \in R_y} f_{X,Y}(x, y),$ for all $x \in R_x$.
Get back original distribution of X.

Marginal distribution of Y sum over all possible values of X:
$f_Y(y) = \sum_{x \in R_x} f_{X,Y}(x, y),$ for all $y \in R_y$

Multiple random variables and joint distributions III

Conditional distribution of X given Y:

$f_{X|Y}(x|y) = \frac{P(X=x, Y=y)}{P(Y=y)} = \frac{f_{X,Y}(x,y)}{f_Y(y)}$

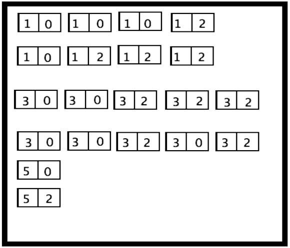
$y = 0 : f_{X|Y}(1|0) = 3/5, f_{X|Y}(3|0) = 2/5, f_{X|Y}(5|0) = 0.$
$y = 2 : ...$
$x = 1 : f_{Y|X}(0|1) = 3/4, f_{Y|X}(2|1) = 1/4.$
$x = 3 : ...$
$x = 5 : ...$

Multiple random variables and joint distributions IV

Independence: $f_{X,Y}(x, y) = f_X(x)f_Y(y)$, for all $x \in R_x, y \in R_y$.
Joint distribution is product of the marginals.
Implies:
Conditional distribution is same as marginal: $f_{X|Y}(x|y) = f_X(x)$.
Knowing value of Y does not change distribution on x.

Multiple random variables and joint distributions V

New Box: same marginal distribution of X and Y but now they are independent:



Fill the joint distribution table, verify that marginals are the same and verify independence:

Y \X	1	3	5	Marg Y
0				
2				
Marg X				

Law of the unconscious statistician I

If $X : S \to R$ is a random variable, and $g : R \to R$ is a function then $Y = g(X)$ is also a random variable with range $R_Y = g(R_X)$.
Example: For the box example from last week take

$$g(x) = \begin{cases} 3 & \text{if } x = 1 \\ 1 & \text{if } x > 1 \end{cases}.$$

Then $Y = g(X)$ has range $R_Y = \{1, 3\}$ and distribution
$f_Y(1) = P(Y = 1) = P(X = 3 \text{ or } X = 5) = 6/10$.
$f_Y(3) = P(Y = 3) = P(X = 1) = 4/10$.
So we can compute the mean: $E(Y) = 1 \cdot 6/10 + 3 \cdot 4/10 = 1.8$

Sometimes it is difficult to compute the new distribution of Y.
It's still easy to compute $E(Y)$ in terms of the distribution of X:
$E(Y) = Eg(X) = \sum_{k=1}^{K} g(x_k) f_X(x_k)$.
So in our example we compute

$$\begin{aligned} E(Y) =& g(1) \cdot 2/5 + g(3) \cdot 1/2 + g(5) \cdot 1/10 \\ =& 3 \cdot 2/5 + 1 \cdot 1/2 + 1 \cdot 1/10 = 1.8 \end{aligned}$$

Extension of law of unconscious statistician I

With two random variables X, Y we can create a new random variable $Z = g(X, Y)$ with $g : R \times R \rightarrow R$.
For example $g(x, y) = x + y$ or $g(x, y) = x \cdot y$.

Again we could try to compute the distribution of Z - hard (conscious statistician).

Or we have:
$Eg(X, Y) = \sum_{x \in R_x, y \in R_y} g(x, y) f_{X,Y}(x, y)$

Try verifying this with the function $g(x, y) = x + y$ and our box of labeled cards.

Extension of law of unconscious statistician II

Consequences of law of unconscious statistician:
$E(X+Y) = EX + EY$.

$$
\begin{aligned}
E(X+Y) &= \sum_{x \in R_x, y \in R_y} (x+y) f_{X,Y}(x,y) \\
&= \sum_{x \in R_x, y \in R_y} x f_{X,Y}(x,y) + \sum_{x \in R_x, y \in R_y} y f_{X,Y}(x,y) \\
\text{Why?} \quad &= \sum_{x \in R_x} x f_X(x) + \sum_{y \in R_y} y f_Y(y) \\
&= EX + EY.
\end{aligned}
$$

More generally, if $X_1, \ldots, X_N$ are random variables:
$E\left(\sum_{n=1}^{N} X_n\right) = \sum_{n=1}^{N} EX_n$.
To know the mean of the sum we don't need to know the joint distribution!

Extension of law of unconscious statistician III

If X, Y are independent: $E(X \cdot Y) = EX \cdot EY$

$$
\begin{aligned}
E(X \cdot Y) &= \sum_{x \in R_x, y \in R_y} x \cdot y \cdot f_{X,Y}(x,y) \\
\text{Independence:} \quad &= \sum_{x \in R_x, y \in R_y} x \cdot y \cdot f_X(x) f_Y(y) \\
\text{Why?} \quad &= \sum_{x \in R_x} x f_X(x) \sum_{y \in R_y} y f_Y(y) \\
&= EX \cdot EY.
\end{aligned}
$$

Extension of law of unconscious statistician IV

Variance of a random variable X with mean μ_X take
$g(x) = (x - \mu_X)^2$.
$Var(X) = Eg(X)$.

Show that:

$Var(X) = EX^2 - \mu_X^2$
$E(aX + b) = aE(X) + b.$
$Var(aX + b) = a^2 Var(X).$

Covariance of two random variables I

Covariance of two random variables X, Y with means μ_X, μ_Y.
Take $g(x, y) = (x - \mu_X) \cdot (y - \mu_Y)$.
$Cov(X, Y) = Eg(X, Y)$ - A measure of how the variables 'covary'.

$Cov(X, Y) > 0 \longrightarrow$ when X increases Y tends to increase.
$Cov(X, Y) < 0 \longrightarrow$ when X increases Y tends to decrease.

Show that $Cov(aX, bY) = a \cdot bCov(X, Y)$.
Changing the units of a measurement will change covariance.

Correlation $\rho(X, Y)$ does not depend on units of measurement:
$\rho(X, Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$
Show that: $\rho(aX, bY) = \rho(X, Y)$.

Covariance of two random variables II

Independence:

Show that $Cov(X, Y) = EXY - \mu_X \mu_Y$.
Conclude: If X, Y are independent $Cov(X, Y) = 0$. (Converse is not true.)

Variance of a sum of random variables I

$$
\begin{aligned}
Var(X+Y) =& E[(X+Y)^2] - [E(X+Y)]^2 \\
=& E[X^2 + 2XY + Y^2] - [E(X)]^2 - 2E(X)E(Y) - [E(Y)]^2 \\
=& E(X^2) - [E(X)]^2 + E(Y^2) - [E(Y)]^2 \\
& + 2E(XY) - 2E(X)E(Y) \\
=& Var(X) + Var(Y) + 2Cov(X,Y)
\end{aligned}
$$

Conclude: If X, Y independent: $Var(X+Y) = Var(X) + Var(Y)$.

More generally, if $X_1, \ldots, X_N$ are independent then

$Var(\sum_{n=1}^{N} X_n) = \sum_{n=1}^{N} Var(X_n)$.

Properties of sample average I

We draw with replacement form a box N times and record the number on each draw as $X_1, X_2, \ldots, X_N$. Because we draw with replacement we can *assume* that the variables X_n are independent. Let μ_B be the *average of the box* and let σ_B^2 be the mean square deviation (MSD) of the box.

Recall that $EX_n = \mu_B$ and $VarX_n = \sigma_B^2$ for each n.
Denote the sample average as $\overline{X} = \frac{1}{N}\sum_{n=1}^{N} X_n$.

$$
\begin{aligned}
E\overline{X} &= \frac{1}{N} E \sum_{n=1}^{N} X_n \\
\text{Why?} \quad &= \frac{1}{N} \sum_{n=1}^{N} EX_n \\
\text{Why?} \quad &= \mu_B .
\end{aligned}
$$

Properties of sample average II

The variance of the sample average

$$
\begin{aligned}
Var(\overline{X}) &= Var(\frac{1}{N}\sum_{n=1}^{N} X_n) \\
&= \frac{1}{N^2} Var(\sum_{n=1}^{N} X_n) \\
\text{Independence: } &= \frac{1}{N^2}\sum_{n=1}^{N} Var(X_n) \\
&= \frac{\sigma_B^2}{N}
\end{aligned}
$$

So the expected value of the average of the sample is the average of the box no matter how large the sample.

The variance of the sample average *decreases* as the sample size N increases $\rightarrow$ The Law of Large Numbers.

Properties of sample average III

Example: $X_1, \ldots, X_N$ are draws from a box of 0's and 1's with replacement. Assume fraction of 1's in the box is p.

Each X_n is Ber(p), i.e. $EX_n = p, Var(X_n) = p(1-p)$.

Since the draws are with replacement we can assume they are independent and so writing $S = \sum_{n=1}^{N} X_n$ we have that S is Binomial(N,p).

$f_S(n) = \binom{N}{n} p^n (1-p)^{N-n}$.

We can compute ES using the definition $\sum_{n=0}^{N} n \binom{N}{n} p^n (1-p)^{N-n}$ but that requires some complicated algebra.

Properties of sample average IV

Instead we use the rules for mean and variance of a sum:
$E(S) = \sum_{n=1}^{N} EX_n = Np.$

And since X_n are independent:
$Var(S) = \sum_{n=1}^{N} Var(X_n) = Np(1-p).$

And for the sample average:
$E\overline{X} = E\frac{S}{N} = \frac{1}{N}ES = p.$

$Var(\overline{X}) = Var(\frac{S}{N}) = \frac{1}{N^2}Np(1-p) = \frac{p(1-p)}{N}.$

The expected value of the sample average is p - the proportion of 1's in the box: it is centered in the 'right' place.
The variance of the sample average decreases with sample size.

Moment generating function I

For any random variable X and for any t we write:

$m_X(t) = Ee^{tX} = \sum_{x \in R_X} e^{tx} f_X(x)$.

Note that the first derivative is:

$m'_X(t) = \sum_{x \in R_X} x e^{tx} f_X(x)$, so that $m'_X(0) = EX$.

The second derivative is:

$m''_X(t) = \sum_{x \in R_X} x^2 e^{tx} f_X(x)$, so that $m''_X(0) = EX^2$

The r'th derivative is

$m^{(r)}_X(t) = \sum_{x \in R_X} x^r e^{tx} f_X(x)$, so that $m^{(r)}_X(0) = EX^r$.

$m_X(t)$ is called the *moment generating function* of X.
Moment generating functions are useful because of the following fact:

Two random variables with the same moment generating function have the same distribution.

Moment generating function II

Examples:
Bernoulli random variable:

$$m(t) = (e^{0 \cdot t} \cdot (1 - p) + e^{1 \cdot t} p) = 1 - p + pe^t.$$
$$EX = m'(0) = p, E(X^2) = m''(0) = p.$$

Moment generating function III

Binomial random variable:

$$
\begin{aligned}
m(t) &= \sum_{n=0}^{N} e^{tn} \binom{N}{n} p^n (1-p)^{N-n} \\
&= \sum_{n=0} \binom{N}{n} (pe^t)^n (1-p)^{N-n} \\
\text{(Binomial formula)} \quad &= [pe^t + (1-p)]^N
\end{aligned}
$$

$m'(t) = N[pe^t + (1-p)]^{N-1} pe^t$

$m''(t) = N(N-1)[pe^t + (1-p)]^{N-2} p^2 e^{2t} + N[pe^t + (1-p)]^{N-1} pe^t.$

$EX = m'(0) = Np.$

$E(X^2) = N(N-1)p^2 + Np = Np(1-p) + (Np)^2.$

$Var(X) = E(X^2) - (EX)^2 = Np(1-p) + (Np)^2 - (Np)^2 = Np(1-p).$