Introduction to Probability

Lecture 3: Distributions

By

Mr. Leonard Thuo

Overview

- Review of basic statistical properties
- Define Random variables
- Distributions

Review of basic statistical concepts

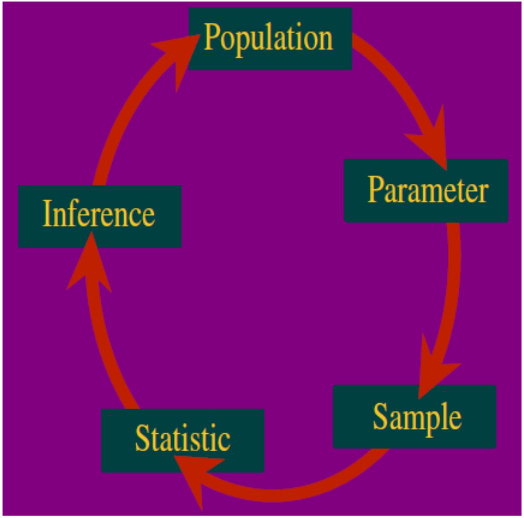
Population Set of measurements or items of interest, e.g.,Kenyan males between the ages of 18-74, intravenous (IV) drug users, smokers, etc. A characteristic of the population is called a parameter

Sample Any subset from the population of interest. A characteristic of the sample is called a statistic

What is inferential statistics

- Interested in a particular characteristic of the population (a parameter).
- To get an idea about the parameter, we select a (random) sample1 and observe a related characteristic in the sample (a statistic).
- Based on assumption of the behavior of this statistic we make guesses about the related population parameter.
- This is called inference, since we infer something about the population.
- Statistical inference is performed in two ways: **Testing of hypotheses and estimation**

Progression of statistical analysis



Population

Parameter

Sample

Statistic

Inference

Random variables and distributions

A random variable: An entity whose observed values are the outcomes of a random experiment. In this sense, their value cannot be a priori determined. That is, we do not know what they are going to be before we collect the sample, run the experiment, etc.

Probability distribution: The mechanism determining the probability or chance of observing each individual value of the random variable (as it literally distributes the probability among all the possible values of the random variables). Probability distributions are defined through frequency tables, graphs, or mathematical expressions.

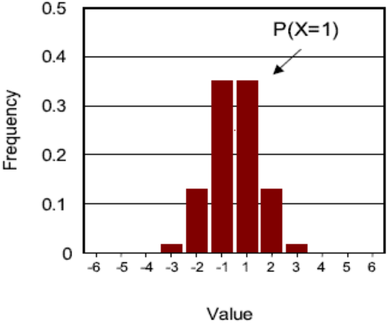
Types of probability distributions

There are two type corresponding to the two kinds of random variables:

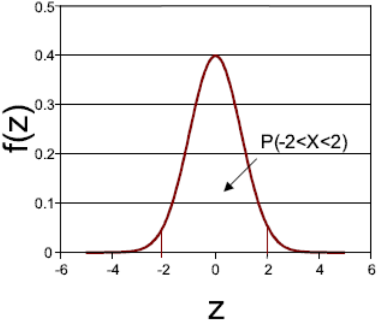
Discrete probability distributions: These specify the chance of observing a small countable number of possible values (e.g., race, sex). Note that large or infinite numbers of countable values are handled by continuous distributions.

Continuous probability distributions: These handle cases were all possible (real) numbers can be observed (e.g., height or weight)

Examples of probability distribution functions



Discrete probability distribution



Continuous probability distribution

Bernoulli distribution B(p)

Let X be a discrete random variable. Let its support be:(0,1). Let p be in the interval (0,1) We say that X has a Bernoulli distribution with parameter p if its probability mass function is:

$$P(x) = \begin{cases} p & \text{if } X = 1 \\ 1 - p & \text{if } X = 0 \\ 0 & \text{otherwise} \end{cases}$$

The mean and variance are

$$E(X) = p$$

$$Var(X) = pq \text{ where } q = (1 - p)$$

Examples Tossing of a coin

Binomial distribution B(n,p)

Describes how a number (n) of binary events (having only two possible outcomes) behave. These events, called Bernoulli trials have the following properties

1. Each event is identical to the others
2. All Bernoulli trial are mutually independent from all the others (i.e., information on the outcome of one does not affect the chances of any other)
3. There are two possible outcomes usually denoted as "success"=1 and "failure"=0
4. The probability of a "success"=π is the same for all trials

The formula producing the probabilities of all possible arrangements of successes and failures is:

$$\begin{pmatrix} n \\ k \end{pmatrix} \pi^k (1-\pi)^{n-k}$$

where $\begin{pmatrix} n \\ k \end{pmatrix}$ number of ways of exactly k successes in n trials

Cont: Binomial

The mean and variance are

$$E(X) = np$$

$$Var(X) = npq$$

Approximate binomial distribution

Example of Binomial distribution

- Ten patients are treated surgically. For ecah person there is a 70% chance of a successful surgery i.e p=0.7
- What is the probability of only 5 successful surgeries?
- Model: the number of successful surgeries in 10 operations is a binomial distributed random variable
- The model is appropriate, since one can assume
    - A surgery is a Bernoulli experiment
    - Whether one surgery is successful is independent of whether the following ones will be successful
    - The probability that a surgery is successful for each person is constant

Estimating the model parameter

$$\hat{p} = 0.70$$

Prediction given by the model

$$P(S_n = k) = P(S_{10} = 5) = \left(\begin{array}{c} 10 \\ 5 \end{array}\right) 0.70^5 (1-0.70)^{10-5} = 0.103$$

Poisson Distribution

- A random variable X with Poisson distribution is useful for characterizing counts, such as the number of occurrences of an event.
- It is characterized by a rate parameter λ. The possible values of X are $0, 1, 2, \ldots$.
- Its mass function is

$$P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}$$

- The rate λ is the average of X
- The mass function sums to one

Types of processes that can be modeled using Poisson distribution

- Number of seizures per week for an epileptic individual
- Number of hospitalizations per month for a chronically ill individual
- Number of occurrences of a rare event over a fixed period (e.g.lottery winners, cases of a rare disease)

Calculations with the Poisson distribution

- Let X denote the number of XDR TB cases in a single month at a clinic in South Africa.
- It follows a Poisson distribution with rate parameter $\lambda = 3$, which means on average, there are 3 cases per month.

Example 1. What is the probability of observing zero cases?

$$P(X = 0) = \frac{e^{-3}3^0}{0!} = 0.0498$$

Example 2. What is the probability of observing 2 cases?

$$P(X = 2) = \frac{e^{-3}3^2}{2!} = 0.2240$$

Poisson as an approximation to the binomial distribution

The Poisson makes a good approximation to the binomial model when the number of trials is very large and the probability of success is small (for example, lottery).

Normal Distribution $N(\mu, \sigma^2)$

$$f_N(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\{\frac{-(x-\mu)^2}{2\sigma^2}\}$$

The Standard Normal Distribution $N(0,1)$

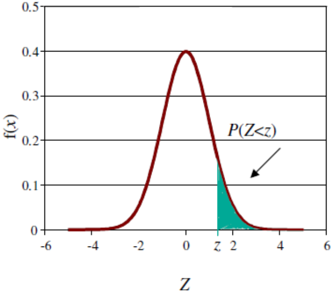
$$f(x) = \frac{1}{\sqrt{2\pi}}\exp\{\frac{-1}{2}x^2\}$$

- Any normal distribution can be transformed to $N(0,1)$ using $Z = \frac{(X-\mu)}{\sigma}$

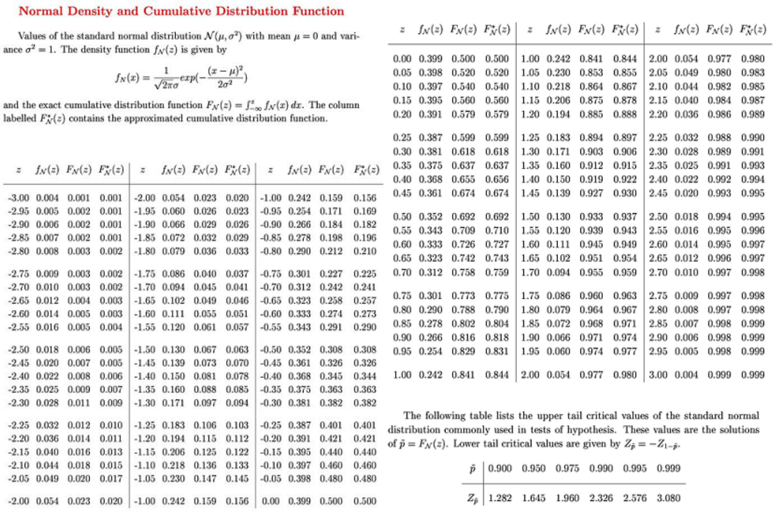
Cont:Normal

- described by two parameters μ and σ
- symmetric
- for a $N(0,1)$ random variable, probability for values larger than 1.96s (2.58s) relative to the mean is smaller than 5% (1%).
- Nearly all values lie within $[-3\sigma, 3\sigma]$.
- It is symmetrical and centered around μ. Each probability is determined as the area between the density curve and the x axis

Probabilities under standard normal distribution



Normal distribution table



Normal Density and Cumulative Distribution Function

Values of the standard normal distribution $\mathcal{N}(\mu, \sigma^2)$ with mean $\mu = 0$ and variance $\sigma^2 = 1$. The density function $f_{\mathcal{N}}(z)$ is given by

$$f_{\mathcal{N}}(x) = \frac{1}{\sqrt{2\pi}\sigma} exp(-\frac{(x-\mu)^2}{2\sigma^2})$$

and the exact cumulative distribution function $F_{\mathcal{N}}(z) = \int_{-\infty}^{z} f_{\mathcal{N}}(x)\,dx$. The column labelled $F^*_{\mathcal{N}}(z)$ contains the approximated cumulative distribution function.

z	$f_{\mathcal{N}}(z)$	$F_{\mathcal{N}}(z)$	$F^*_{\mathcal{N}}(z)$	z	$f_{\mathcal{N}}(z)$	$F_{\mathcal{N}}(z)$	$F^*_{\mathcal{N}}(z)$	z	$f_{\mathcal{N}}(z)$	$F_{\mathcal{N}}(z)$	$F^*_{\mathcal{N}}(z)$
-3.00	0.004	0.001	0.001	-2.00	0.054	0.023	0.020	-1.00	0.242	0.159	0.156
-2.95	0.005	0.002	0.001	-1.95	0.060	0.026	0.023	-0.95	0.254	0.171	0.169
-2.90	0.006	0.002	0.001	-1.90	0.066	0.029	0.026	-0.90	0.266	0.184	0.182
-2.85	0.007	0.002	0.001	-1.85	0.072	0.032	0.029	-0.85	0.278	0.198	0.196
-2.80	0.008	0.003	0.002	-1.80	0.079	0.036	0.033	-0.80	0.290	0.212	0.210
-2.75	0.009	0.003	0.002	-1.75	0.086	0.040	0.037	-0.75	0.301	0.227	0.225
-2.70	0.010	0.003	0.002	-1.70	0.094	0.045	0.041	-0.70	0.312	0.242	0.241
-2.65	0.012	0.004	0.003	-1.65	0.102	0.049	0.046	-0.65	0.323	0.258	0.257
-2.60	0.014	0.005	0.003	-1.60	0.111	0.055	0.051	-0.60	0.333	0.274	0.273
-2.55	0.016	0.005	0.004	-1.55	0.120	0.061	0.057	-0.55	0.343	0.291	0.290
-2.50	0.018	0.006	0.005	-1.50	0.130	0.067	0.063	-0.50	0.352	0.308	0.308
-2.45	0.020	0.007	0.005	-1.45	0.139	0.073	0.070	-0.45	0.361	0.326	0.326
-2.40	0.022	0.008	0.006	-1.40	0.150	0.081	0.078	-0.40	0.368	0.345	0.344
-2.35	0.025	0.009	0.007	-1.35	0.160	0.088	0.085	-0.35	0.375	0.363	0.363
-2.30	0.028	0.011	0.009	-1.30	0.171	0.097	0.094	-0.30	0.381	0.382	0.382
-2.25	0.032	0.012	0.010	-1.25	0.183	0.106	0.103	-0.25	0.387	0.401	0.401
-2.20	0.036	0.014	0.011	-1.20	0.194	0.115	0.112	-0.20	0.391	0.421	0.421
-2.15	0.040	0.016	0.013	-1.15	0.206	0.125	0.122	-0.15	0.395	0.440	0.440
-2.10	0.044	0.018	0.015	-1.10	0.218	0.136	0.133	-0.10	0.397	0.460	0.460
-2.05	0.049	0.020	0.017	-1.05	0.230	0.147	0.145	-0.05	0.398	0.480	0.480
-2.00	0.054	0.023	0.020	-1.00	0.242	0.159	0.156	0.00	0.399	0.500	0.500

z	$f_{\mathcal{N}}(z)$	$F_{\mathcal{N}}(z)$	$F^*_{\mathcal{N}}(z)$	z	$f_{\mathcal{N}}(z)$	$F_{\mathcal{N}}(z)$	$F^*_{\mathcal{N}}(z)$	z	$f_{\mathcal{N}}(z)$	$F_{\mathcal{N}}(z)$	$F^*_{\mathcal{N}}(z)$
0.00	0.399	0.500	0.500	1.00	0.242	0.841	0.844	2.00	0.054	0.977	0.980
0.05	0.398	0.520	0.520	1.05	0.230	0.853	0.855	2.05	0.049	0.980	0.983
0.10	0.397	0.540	0.540	1.10	0.218	0.864	0.867	2.10	0.044	0.982	0.985
0.15	0.395	0.560	0.560	1.15	0.206	0.875	0.878	2.15	0.040	0.984	0.987
0.20	0.391	0.579	0.579	1.20	0.194	0.885	0.888	2.20	0.036	0.986	0.989
0.25	0.387	0.599	0.599	1.25	0.183	0.894	0.897	2.25	0.032	0.988	0.990
0.30	0.381	0.618	0.618	1.30	0.171	0.903	0.906	2.30	0.028	0.989	0.991
0.35	0.375	0.637	0.637	1.35	0.160	0.912	0.915	2.35	0.025	0.991	0.993
0.40	0.368	0.655	0.656	1.40	0.150	0.919	0.922	2.40	0.022	0.992	0.994
0.45	0.361	0.674	0.674	1.45	0.139	0.927	0.930	2.45	0.020	0.993	0.995
0.50	0.352	0.692	0.692	1.50	0.130	0.933	0.937	2.50	0.018	0.994	0.995
0.55	0.343	0.709	0.710	1.55	0.120	0.939	0.943	2.55	0.016	0.995	0.996
0.60	0.333	0.726	0.727	1.60	0.111	0.945	0.949	2.60	0.014	0.995	0.997
0.65	0.323	0.742	0.743	1.65	0.102	0.951	0.954	2.65	0.012	0.996	0.997
0.70	0.312	0.758	0.759	1.70	0.094	0.955	0.959	2.70	0.010	0.997	0.998
0.75	0.301	0.773	0.775	1.75	0.086	0.960	0.963	2.75	0.009	0.997	0.998
0.80	0.290	0.788	0.790	1.80	0.079	0.964	0.967	2.80	0.008	0.997	0.998
0.85	0.278	0.802	0.804	1.85	0.072	0.968	0.971	2.85	0.007	0.998	0.999
0.90	0.266	0.816	0.818	1.90	0.066	0.971	0.974	2.90	0.006	0.998	0.999
0.95	0.254	0.829	0.831	1.95	0.060	0.974	0.977	2.95	0.005	0.998	0.999
1.00	0.242	0.841	0.844	2.00	0.054	0.977	0.980	3.00	0.004	0.999	0.999

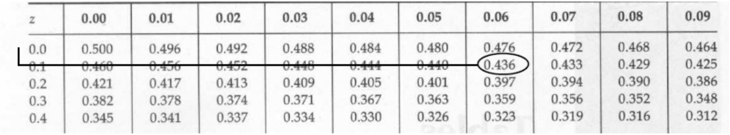
The following table lists the upper tail critical values of the standard normal distribution commonly used in tests of hypothesis. These values are the solutions of $\tilde{p} = F_{\mathcal{N}}(z)$. Lower tail critical values are given by $Z_{\tilde{p}} = -Z_{1-\tilde{p}}$.

$\tilde{p}$	0.900	0.950	0.975	0.990	0.995	0.999
$Z_{\tilde{p}}$	1.282	1.645	1.960	2.326	2.576	3.080

Reading Normal distribution table

To find a probability $P(Z > z)$ in the standard normal table from the above Table

- we search for z by proceeding down the left margin of the table going to a row that is just below z.
- Then we go across to a column that is as closest to z.
- The following figure helps clarify this for the case $P(Z > 0.16)$.



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.500	0.496	0.492	0.488	0.484	0.480	0.476	0.472	0.468	0.464
0.1	0.460	0.456	0.452	0.448	0.444	0.440	0.436	0.433	0.429	0.425
0.2	0.421	0.417	0.413	0.409	0.405	0.401	0.397	0.394	0.390	0.386
0.3	0.382	0.378	0.374	0.371	0.367	0.363	0.359	0.356	0.352	0.348
0.4	0.345	0.341	0.337	0.334	0.330	0.326	0.323	0.319	0.316	0.312

This means that $P(Z > 0 : 16) = 0.436$

Reading normal dist table

Take advantage of the following features of normal distribution when reading table

1. The symmetry of the standard normal curve around zero (its mean). Thus, $P(Z \geq z) = P(Z \leq -z)$, where $z \geq 0$.
2. As in any distribution the area under the curve is equal to 1. Thus, two complementary events, $P(Z \geq z) = 1 - P(Z \leq z)$.

Tips reading normal dist table

We are usually faced with two problems:

1. Given a number $z > 0$ (say) find p such that the following is true:
   1.1 $P(Z \geq z) = p$. To do this we read p directly from standard normal table
   1.2 $P(Z \leq -z) = p$. In this case, we read $p_1 = P(Z \geq z)$ from the normal table, which by the symmetry of the normal distribution is equal to p
   1.3 $P(Z \leq z) = p$. We read $p_1 = P(Z \geq z)$ from the normal table. Now $p = 1 - p_1$ since $P(Z \leq z)$ and $P(Z \geq z)$ are complementary events

Tips reading normal dist table

1. $P(Z \geq -z) = p$. We Read $p_1 = P(Z \geq z)$ from the normal table and then $p = 1 - p_1$
2. Assuming that $z_1 \leq z_2$ we want to calculate $P(z_1 \leq Z \leq z_2) = p$. Since this is the area below z_2) i.e., $P(Z \leq z_2)$ with the "piece" $P(Z \leq z_1)$ removed, this is $P(z_1 \leq Z \leq z_2) = P(Z \leq z_2) - P(Z \leq z_1)$ (see above for the manner that these probabilities are calculated). In the special case $P(z \leq Z \leq z) = 1 - 2P(Z > z)$

Tips reading normal dist table

Given a probability p and z such that the following is true

1. $P(Z \geq z) = p$ If $p = 0 \geq 5$. Then $z \geq 0$ and we look up p in the table. On the other hand, if $p \leq 0.5$ then $z \leq 0$ and we look up $p_1 = 1 - p$ in the table. z is the negative of the number located in the table
2. $P(Z \leq z) = p$ If $p \leq 0.5$ then $z \leq 0$ and again we look up p in table. z is the negative of the number located there. On the other hand, if $p \geq 0.5$ then $z \geq 0$ and we look up $p_1 = 1 - p$ in the table.
3. $P(-z \leq Z \leq z) = p$. Look up $p_1 = (1 - p)/2$ in the table. z is the closest number while -z is its negative.