Probability Outline for STAT112

Charles Fleming

January 10, 2018

Theorem 1. *If E_1 and E_2 are subsets of Ω and if they are disjoint, then $P(E_1 \cup E_2) = P(E_1) + P(E_2)$.*

Sometimes, when working with sets, our attention is focused on only one of them while the rest are put together into the complement.

Definition 1. *The **complement** of E, denoted by E^c, is the set of all elements of Ω which are not elements of E.*

Corollary 1. *$E \cup E^c = \Omega$.*

Theorem 2. *$P(E^c) = 1 - P(E)$.*

Theorem 3. *If G_1 and G_2 are subsets of Ω, then $P(G_1 \cup G_2) = P(G_1) + P(G_2) - P(G_1 \cap G_2)$.*

1 Axioms of Probability

Axiom 1. *$0 \leqslant P(E) \leqslant 1$.*

Axiom 2. *$P(\Omega) = 1$.*

Axiom 3. *If $E_1, E_2, \cdots, E_n$ are pairwise disjoint subsets of Ω, then $P(E_1 \cup E_2 \cup \cdots \cup E_n) = P(E_1) + P(E_2) + \cdots + P(E_n)$.*

Definition 2. *A function, P, that satisfies Axioms 1–3 is called a **probability**.*

Definition 3. *Denote the **conditional probability** by $P(A|B)$ which means: the probability of the event, A, given that the event, B, has occurred. Also $P(A|B) = \frac{P(A \cap B)}{P(B)}$.*

Definition 4. *If $P(A|B) = P(A)$, then the events A and B are said to be **independent**.*

Theorem 4. *Suppose that events A and B are independent, then* $P(A \cap B) = P(A)P(B)$.

Theorem 5. *Let A and E be events, then* $P(A) = P(A|E)P(E) + P(A|E^c)P(E^c)$.

Definition 5. *A function, X, which maps an outcome of the sample space to a number on the real line is called a* ***random variable****.*

Definition 6. *The* ***expected value*** *of a random variable, X, is defined to be:*

$$E[X] = \sum_{\substack{all\ possible \\ values\ of\ X}} kP(X = k).$$

(The summation $\sum$ *is taken over all possible values of X)*

Definition 7. *The* ***variance*** *of a random variable, X, is defined to be:*

$$var(X) = \sum_{\substack{all\ possible \\ values\ of\ X}} (k - E[X])^2 P(X = k).$$

(The summation is taken over all possible values of X)

Definition 8. *The set of values* $\{P(X = 0),\ P(X = 1),\ \ldots,\ P(X = n)\}$ *is called the* ***probability distribution*** *or the* ***probability mass function*** *of the random variable X.*

Observations 1.

1. $P(X = i) \geqslant 0$.
2. $\sum_{\substack{all\ possible \\ values\ of\ X}} P(X = k) = 1$.
3. $P(X \leqslant k) = P(X = 0) + P(X = 1) + \ \ldots\ + P(X = k - 1) + P(X = k)$.
4. $P(X > k) = 1 - P(X \leqslant k)$.
5. $P(X = k) = P(X \leqslant k) - P(X \leqslant k - 1)$.

Only two outcomes characterize the Bernoulli distribution.

The Uniform distribution is characterized by the fact that for all values of the random variable, X, the probabilities are the same.

Bernoulli b(1,p)	Discrete Uniform (a,b)	Binomial b(n,p)
$E[X] = p$	$E[X] = \frac{a+b}{2}$	$E[X] = np$
$var(X) = pq$	$var(X) = \frac{(b-a)(b-a+2)}{12}$	$var(X) = npq$

The Binomial distribution is characterized by n independent trials for which each trial has two outcomes with a probability of success, p. If $X \sim b(n,p)$, then

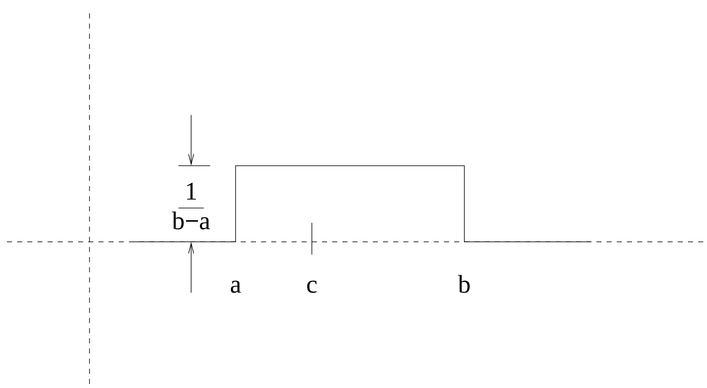
$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \quad \text{for} \quad k = 0,\ 1,\ 2,\ \ldots,\ n \tag{1}$$

Definition 9. *The symbol $\binom{n}{k}$ is the number of possible subsets of size k that can be drawn from a set of n objects.*

Theorem 6.

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \quad \textit{where } n!=n(n\text{-}1)(n\text{-}2)\ldots(3)(2)(1)$$

The best way to describe the continuous Uniform distribution is with a picture of its probability density function as shown here.



The equation of the probability density function (pdf) of the Normal distribution, $N(\mu, \sigma^2)$, is: $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$.

Contiuous Uniform U(a,b)
$E[X] = \frac{a+b}{2}$
$var(X) = \frac{(b-a)^2}{12}$

Normal N(μ,σ^2)
$E[X] = \mu$
$var(X) = \sigma^2$

An important property of the Normal distribution is: If $X \sim N(\mu, \sigma^2)$, then $aX + b \sim N(a\mu + b, a^2\sigma^2)$ for the two arbitrary constants, a and b. In particular, let $a = \frac{1}{\sigma}$ and $b = -\frac{\mu}{\sigma}$, then $aX + b = \frac{X-\mu}{\sigma}$; moreover, $a\mu + b = 0$, and $a^2\sigma^2 = 1$, hence $\frac{X-\mu}{\sigma} \sim N(0,1)$.

Definition 10. $\frac{X-\mu}{\sigma}$ *is called the* ***population z-score*** *of* X. $\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$ *is the* ***population z-score*** *of* $\bar{X}$.

Definition 11. $\frac{x_i-\bar{x}}{s}$ *is called the* ***sample z-score*** *of* x_i.

Note: $\frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}} \sim T_{n-1}$ where T_{n-1} is Student's T distribution with n-1 degrees of freedom and $E[X_i] = \mu \quad \forall i$.

Definition 12. *i.i.d. means* ***independent identically distribution***.

lemma 1. *If* $X_1,\ X_2,\ \ldots,\ X_n$ *are i.i.d., then* $E[X_1] = E[X_2] = \ldots = E[X_n]$ *and* $var(X_1) = var(X_2) = \ldots = var(X_n)$.

Theorem 7. *If X and Y are random variables, then E[X+Y]=E[X]+E[Y].*

Theorem 8. *If a and b are constants, then E[aX+b]=aE[X]+b and var(aX+b)=*$a^2 var(X)$.

Theorem 9. *If X and Y are independent random variables, then var(X+Y)=var(X)+var(Y).*

Theorem 10. *If* $X_1,\ X_2,\ \ldots,\ X_n$ *are i.i.d. each with mean* μ *and variance* σ^2, *and* $\bar{x} = \frac{X_1+\cdots+X_n}{n}$, *then*

$$E[\bar{x}] = \mu \text{ and } var(\bar{x}) = \frac{\sigma^2}{n}$$

Note: for a finite population $\widehat{var(\bar{x})} = \left(\frac{N-n}{N}\right)\frac{s^2}{n}$. *See Theory of Survey Sampling for STAT112*

Definition 13. *If* $\hat{\Theta}$ *is an estimator of* Θ, *then* $\hat{\Theta}$ *is called an* ***unbiased estimator*** *of* Θ, *if* $E[\hat{\Theta}] = \Theta$.

If X_i are i.i.d. $N(\mu, \sigma^2)$ and $\bar{x} = \frac{X_1+X_2+\cdots+X_n}{n}$, then $E[\bar{x}] = \mu$; hence $\bar{x}$ is called an unbiased estimator of μ. Also, $E[s^2] = \sigma^2$; hence s^2 is called an unbiased estimator of σ^2 where $s^2 = \frac{\sum_{i\in\mathcal{S}}(X_i-\bar{X})^2}{n-1}$.

Theorem 11. *(Central Limit Theorem) If* $X_1,\ X_2,\ \ldots\ , X_n$ *are i.i.d. and* $\bar{x} = \frac{X_1+\cdots+X_n}{n}$ *then*

$$\frac{\bar{x} - E[\bar{x}]}{\sqrt{var(\bar{x})}} = \text{sample z-score} \quad \rightarrow \quad N(0,\ 1) \quad \text{as } n \rightarrow \infty$$

Theorem 12. *Given random variables, X and Y, then:*

$$\begin{aligned} E[Y] &= E[E[Y|X]] \\ var(Y) &= var(E[Y|X]) + E[var(Y|X)] \end{aligned}$$