MTH 7241 Fall 2022: Prof. C. King

Notes 3: Sequences and limit theorems

IID sequence

The random variables $Y_1, Y_2, \ldots$ are *independent and identically distributed* or **IID** if they are all independent and have the same pdf. Independence means that

$$\mathbb{P}(Y_{i_1} = k_{i_1}, \ldots, Y_{i_n} = k_{i_n}) = \mathbb{P}(Y_{i_1} = k_{i_1}) \ldots \mathbb{P}(Y_{i_n} = k_{i_n})$$

for all subsets of variables and all values $k_{i_1}, \ldots, k_{i_n}$. In particular they all have the same mean μ and variance σ^2. IID variables arise in sampling problems, where successive independent measurements are made on a random system. For example, Y_n could be the value of the nth roll of a die, or the nth toss of a coin.

$Y_1 \sim U[0,1]$

$Y_2 \sim U[0,1]$ $\quad E[Y_1] = E[Y_2] = 1/2 = \mu$

$VAR[Y_1] = VAR[Y_2] = 1/12 = \sigma^2$

$\vdots$

$Y_n \sim U[0,1]$

$Y_1, Y_2, \ldots Y_n$ are IID $\to$ same pdf, independent

$\Downarrow$ IID

$P(Y_1 \le 1/2, Y_2 \le 3/4) = P(Y_1 \le 1/2)\, P(Y_2 \le 3/4) = (1/2)(3/4)$

Random binary string

Let $Y_1, \ldots$ be IID Bernoulli r.v.'s, with

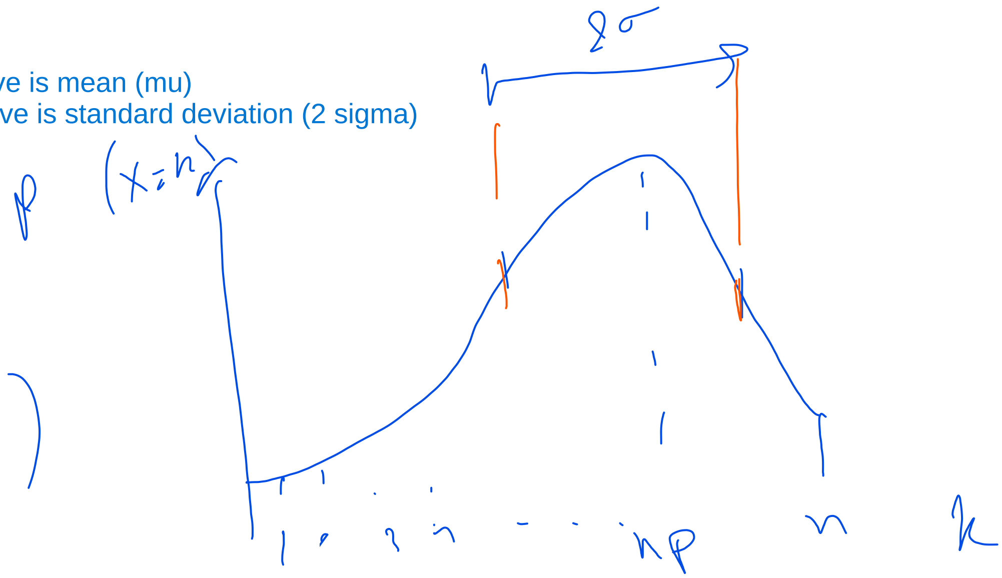
$$\mathbb{P}(Y = 1) = p, \quad \mathbb{P}(Y = 0) = 1 - p$$

The sequence $(Y_1, Y_2, \ldots, Y_n)$ is a random binary n-string, or random bit-string. The sum $\sum_i Y_i$ is the number of 1's in the string, and its pdf is binomial with parameters (n, p).

Peak of bell curve is mean (mu)
Width of bell curve is standard deviation (2 sigma)

$\mu = np$

$\sigma^2 = np(p-1)$



Law of Large Numbers

Let $Y_1, Y_2, \ldots$ be any collection of IID r.v.'s (independent and identically distributed random variables) with common mean μ and variance σ^2. The sample mean of the first n variables is

$$\overline{Y}_n = \frac{1}{n}(Y_1 + \cdots + Y_n)$$

The LNN (Law of Large Numbers) says that $\overline{Y}_n$ converges to the true mean μ as $n \to \infty$.

Note that $\overline{Y}_n = \frac{1}{n}(Y_1 + \cdots + Y_n)$ is the long-run average value of a sequence of measurements. For each possible value y_k define

$$n(k) = \#\{i \, : \, Y_i = y_k\}$$

so then

$$Y_1 + \cdots + Y_n = \sum_k y_k \, n(k)$$

The LLN is equivalent to the statement that the relative frequency of occurrence of each value converges to its probability:

$$\frac{n(k)}{n} \to \mathbb{P}(Y = y_k) \quad \text{as } n \to \infty$$

and therefore

$$\frac{1}{n}(Y_1 + \cdots + Y_n) = \sum_k y_k \, \frac{n(k)}{n} \to \sum_k y_k \, \mathbb{P}(Y = y_k) = \mathbb{E}[Y]$$

Real issue is what 'convergence' means. Given any $\epsilon > 0$, $\delta > 0$, there is $N < \infty$ such that for all $n \geq N$, $\overline{Y}_n$ will with probability at least $1 - \delta$ lie inside the interval $\mu \pm \epsilon$. Or more succinctly, for every $\epsilon > 0$,

$$\lim_{n \to \infty} \mathbb{P}\left(\left| \frac{1}{n}(Y_1 + \cdots + Y_n) - \mu \right| > \epsilon \right) = 0$$

Markov's Inequality

Many useful results in probability are proved using inequalities. The relevant one for the LLN is called Markov's inequality and is easily stated. For any random variable X and for any numbers $a > 0$ and $k > 0$,

$$\mathbb{P}(|X| \geq a) \leq \frac{1}{a^k}\,\mathbb{E}[|X|^k]$$

The proof is easy:

$$\begin{aligned}
\mathbb{E}[|X|^k] &= \sum_i |x_i|^k P(X = x_i) \\
&\geq \sum_{i:|x_i|\geq a} |x_i|^k P(X = x_i) \\
&\geq a^k \sum_{i:|x_i|\geq a} P(X = x_i) \\
&= a^k P(|X| \geq a)
\end{aligned}$$

An important special case of Markov's inequality is called *Chebyshev's inequality*: take $X = Y - \mathbb{E}Y$ and $k = 2$ to get

$$P(|Y - \mathbb{E}Y| \geq a) \leq \frac{1}{a^2} \operatorname{Var}(Y)$$

The LLN follows easily from this. Take

$$X = \frac{1}{n}(Y_1 + \cdots + Y_n)$$

then $\mathbb{E}[X] = \mu$ and $\mathrm{VAR}[X] = \sigma^2/n$, hence

$$P(|X - \mu| \geq \epsilon) \leq \frac{\sigma^2}{n\,\epsilon^2} \to 0$$

Example 1 *Toss an unbiased coin* 1000 *times. The running average* $\overline{Y}_n$ *converges to* 0.5. *Use Markov's inequality to lower bound*

$$P(0.4 < \overline{Y}_n < 0.6) \quad \textit{for } n = 1000$$

Central Limit Theorem

Theorem 1 *Let* $Y_1, Y_2, \dots$ *be IID with finite mean* $\mathbb{E}Y_i = \mu$ *and finite variance* $\mathrm{VAR}[Y_i] = \sigma^2$. *Define*

$$Z_n = \frac{Y_1 + \cdots + Y_n - n\mu}{\sigma\sqrt{n}}$$

Then for all $a < b$, *as* $n \to \infty$,

$$P(a < Z_n \le b) \to P(a < Z \le b) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx$$

The integrand on the right side is the pdf of the standard normal. So another way to state the CLT is

$$Z_n \to Z \sim N(0,1) \quad \text{(convergence in distribution)}$$

Even more informally, we can say that for n large,

$$\bar{Y}_n = \frac{1}{n}\sum_{i=1}^{n} Y_i = \mu + \frac{\sigma}{\sqrt{n}} Z + \cdots$$

or even

$$\sum_{i=1}^{n} Y_i = n\,\mu + \sigma\,\sqrt{n}\, Z + \cdots$$

where $\cdots$ goes to zero faster than the leading order terms as $n \to \infty$. This is the most useful way to to think about the meaning of the CLT, and will guide you in its application.

Example 2 *As noted, the standard normal* $Z \sim N(0,1)$ *has mean* 0 *and variance* 1*, and pdf*

$$f_Z(x) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2}{2}}$$

The general normal $X \sim N(\mu, \sigma^2)$ *has mean* μ *and variance* σ^2*, and its pdf is*

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

The pdf $f_X(x)$ *is used to compute expected values by*

$$\mathbb{E}[X^k] = \int_{-\infty}^{\infty} x^k\, f_X(x)\, dx$$

We note that every normal can be written as a translated and rescaled version of the standard normal: if $X \sim N(\mu, \sigma^2)$ *then*

$$X = \mu + \sigma\, Z$$

meaning that $(X - \mu)/\sigma$ *is a standard normal. This is very useful for calculations, for example: supposing* $X \sim N(2, 9)$*, what is the probability for* X *to be greater than* 8*? We have*

$$X = 2 + 3Z \Rightarrow \mathbb{P}(X > 8) = \mathbb{P}(Z > 2) = 0.023$$

(the number 0.023 *is obtained from z-tables).*

Tables of the Normal Distribution

Probability Content from -oo to Z

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

Far Right Tail Probabilities

Z	P{Z to oo}	Z	P{Z to oo}	Z	P{Z to oo}	Z	P{Z to oo}
2.0	0.02275	3.0	0.001350	4.0	0.00003167	5.0	2.867 E-7
2.1	0.01786	3.1	0.0009676	4.1	0.00002066	5.5	1.899 E-8
2.2	0.01390	3.2	0.0006871	4.2	0.00001335	6.0	9.866 E-10
2.3	0.01072	3.3	0.0004834	4.3	0.00000854	6.5	4.016 E-11
2.4	0.00820	3.4	0.0003369	4.4	0.000005413	7.0	1.280 E-12
2.5	0.00621	3.5	0.0002326	4.5	0.000003398	7.5	3.191 E-14
2.6	0.004661	3.6	0.0001591	4.6	0.000002112	8.0	6.221 E-16
2.7	0.003467	3.7	0.0001078	4.7	0.000001300	8.5	9.480 E-18
2.8	0.002555	3.8	0.00007235	4.8	7.933 E-7	9.0	1.129 E-19
2.9	0.001866	3.9	0.00004810	4.9	4.792 E-7	9.5	1.049 E-21

Example 3 *A fair die is rolled* 1000 *times. Use the CLT to estimate the probability that the number of* 6*'s is greater than* 180*. Define for all* $1 \le i \le 1000$

$$Y_i = \begin{cases} 1 & \textit{if the ith roll is a } 6 \\ 0 & \textit{if the ith roll is not } 6 \end{cases}$$

Let $X = Y_1 + Y_2 + \cdots + Y_{1000}$*, then* X *is equal to the number of* 6*'s rolled. Now* $\mathbb{E}[Y] = 1/6$ *and* $\mathrm{VAR}[Y] = 5/36$*, so the CLT gives*

$$X \simeq \frac{1000}{6} + \sqrt{1000}\,\sqrt{5/36}\,Z$$

Therefore

$$\mathbb{P}(X > 180) \simeq \mathbb{P}\left(Z > \frac{180 - 1000/6}{\sqrt{5000/36}}\right) = \mathbb{P}(Z > 1.13) = 0.13$$

What is the accuracy of the CLT? The Berry-Esseen Theorem gives the general bound

$$\left| P(Z_n < a) - P(Z < a) \right| \leq \frac{C\rho}{\sigma^3 \sqrt{n}}$$

where C is a constant (not more than 0.48), and $\rho = \mathbb{E}|Y_i - \mu|^3$.

Central Limit Theorem: sketch of proof

The main idea of the proof is to consider the moment generating function of Z_n, that is

$$M_n(t) = \mathbb{E}[e^{tZ_n}], \quad t \in \mathbb{R}$$

and show that for every number t it converges to the moment generating function of the standard normal, which is

$$M(t) = \mathbb{E}[e^{tZ}] = e^{t^2/2}$$

(this result for the standard normal follows from a simple calculation using Gaussian integrals). Once this convergence has been shown for every t, it follows that all the moments of Z_n converge to the corresponding moments of Z, and this gives the result. The proof of convergence follows a calculation which we sketch below. First, for each $i = 1, \ldots, n$ we define

$$X_i = \frac{Y_i - \mu}{\sigma}$$

It follows that $\mathbb{E}[X_i] = 0$ and $\mathrm{VAR}[X_i] = 1$, and

$$Z_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i$$

Also, since the $\{X_i\}$ are independent we can compute

$$M_n(t) = \mathbb{E}[e^{tn^{-1/2} \sum_i X_i}] = \Big(\mathbb{E}[e^{tn^{-1/2} X}]\Big)^n$$

Now we just observe that for large n, using the Taylor series for the exponential

$$E[e^{tn^{-1/2}X}] = 1 + E[tn^{-1/2}X] + \frac{1}{2}\mathbb{E}[t^2 n^{-1} X^2] + \cdots = 1 + 0 + \frac{t^2}{2n} + \cdots$$

The remainder terms go to zero faster than n^{-1} as $n \to \infty$, so we can ignore these to compute

$$M_n(t) = \left(1 + \frac{t^2}{2n}\right)^n \to e^{t^2/2} = M(t)$$

Simple Random Walk

The simple random walk starts at zero, and at each step moves one unit either forwards (up) or backwards (down). We write X_n for the position of the random walk after n steps. So the possible values of X_n are

$$Ran(X) = \{-n, -n+2, \ldots, n-2, n\}$$

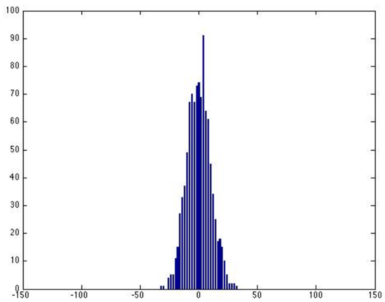
The pdf of X_n is the list of probabilities

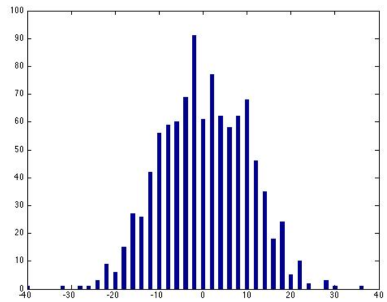
$$\mathbb{P}(X = n - 2k), \quad k = 0, 1, \ldots, n$$

These can be calculated using the binomial distribution as

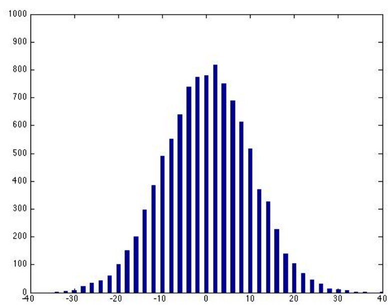
$$\mathbb{P}(X = n - 2k) = \mathbb{P}(k \text{ down}, \, n - k \text{ up}) = \binom{n}{k} 2^{-n}$$

But first let's do an experiment. Repeat the random walk many times and compute the long-run fraction of occurrences of each possible value of X_n. This is called the empirical pdf of X_n. Here $n = 100$ steps and sample 1000 times.





Note that the re-scaled pdf has a definite shape. Here is a longer run, where $n = 100$ and we sample $10,000$ times.



The pdf looks like a normal curve. We can explain this observation using the CLT. The value $X_n = n - 2k$ is achieved by k down steps and $n - k$ up steps. So we can write

$$X_n = S_1 + S_2 + \cdots + S_n$$

where

$$S_k = \begin{cases} 1 & \text{if } k^{th} \text{ step goes up} \\ -1 & \text{if } k^{th} \text{ step goes down} \end{cases}$$

Each step S_k is an independent random variable with a very simple pmf:

$$\mathbb{P}(S_k = 1) = \mathbb{P}(S_k = -1) = \frac{1}{2}$$

Easily calculate that

$$\mathbb{E}[S_k] = 0, \quad \text{VAR}[S_k] = 1$$

Therefore since $X_n = S_1 + \cdots + S_n$ we get

$$\mathbb{E}[X_n] = 0, \quad \text{VAR}[X_n] = n$$

Note that the S_k are IID, so we can apply the CLT. In this case

$$Z_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} S_i = n^{-1/2} X_n$$

and therefore the CLT says

$$Z_n \to Z, \quad X_n \simeq n^{1/2}\, Z$$

So for large n, the position X_n is a rescaled standard normal. This explains the graphs shown above.

χ^2 distribution (Chi-Square)

Normal random variables are important because combinations of them arise in many statistical tests. Suppose that $Z_1, \ldots, Z_m$ are IID standard normal random variables, and suppose that

$$X = Z_1^2 + \cdots + Z_m^2$$

Then the random variable X is said to have a *chi-square* χ^2 *distribution with* m *degrees of freedom.* We indicate this by

$$X \sim \chi^2(m), \quad m = df = \text{degrees of freedom}$$

Properties of chi-square: it is a sum if independent r.v.'s, so

$$\begin{aligned}
\mathbb{E}[X] &= m\,\mathbb{E}[Z^2] = m \\
\mathrm{VAR}[X] &= m\,\mathrm{VAR}[Z^2] = 2m \\
M_X(t) &= (M_{Z^2}(t))^m = (1-2t)^{-m/2} \quad \text{for } t < 1/2
\end{aligned}$$

It is also clear that a sum of independent chi-squares is again chi-square:

$$X \sim \chi^2(m),\ Y \sim \chi^2(n) \Rightarrow X + Y \sim \chi^2(m+n)$$

Use tables (or calculator) to compute probabilities for chi-square variables.

The multinomial distribution is the joint pdf for the frequencies of different results in a sequence of random trials. For example, suppose that a die is rolled ten times, and the following outcomes recorded:

Outcome	1	2	3	4	5	6
Probability	1/6	1/6	1/6	1/6	1/6	1/6
Observed frequency	2	1	0	3	1	3

What is the probability of this result? The answer is

$$\frac{10!}{2!\,1!\,0!\,3!\,1!\,3!}\,(1/6)^{10} = 8.34 \times 10^{-4}$$

Here is the general result: suppose X takes values $x_1, \ldots, x_m$ with probabilities $p_1, \ldots, p_m$. Suppose X is measured N times, and let $R_1, \ldots, R_m$ be the numbers of times each value is measured, so $R_1 + \cdots + R_m = N$. Then the probability to get the frequencies $n_1, \ldots, n_m$ is

$$\mathbb{P}\,(R_1 = n_1, \ldots, R_m = n_m) = \frac{N!}{n_1! \,\cdots\, n_m!}\, p_1^{n_1} \;\ldots\; p_m^{n_m}$$

Goodness of fit distribution

The expected frequency for the ith value is $\mathbb{E}[R_i] = N\,p_i$. The goodness-of-fit random variable is defined as

$$D = \sum_{i=1}^{m} \frac{(R_i - N\,p_i)^2}{N\,p_i} = \sum_{i=1}^{m} \frac{(\text{Observed}_i - \text{Expected}_i)^2}{\text{Expected}_i}$$

The main result is that D has a χ^2 distribution with $m-1$ degrees of freedom. That is,

$$D \sim \chi^2(m-1)$$

(The precise statement says that this holds true in the limit where $N \to \infty$.)

Pearson's Goodness of fit Test

Suppose we have the sequence of observed frequencies $n_1, \ldots, n_m$ from N measurements, and we want to test the hypothesis that the measurements arise from the multinomial distribution with probabilities $p_1, \ldots, p_m$. We use D to design a statistical test based on the observed data.

Null hypothesis H_0: $R_1, \ldots, R_m$ are multinomial $p_1, \ldots, p_m$

Alternative hypothesis H_1: $R_1, \ldots, R_m$ are not multinomial $p_1, \ldots, p_m$.

We compute the test statistic using the observed data and the null hypothesis parameters:

$$d = \sum_{i=1}^{m} \frac{(\text{Observed}_i - \text{Expected}_i)^2}{\text{Expected}_i} = \sum_{i=1}^{m} \frac{(n_i - N\,p_i)^2}{N\,p_i}$$

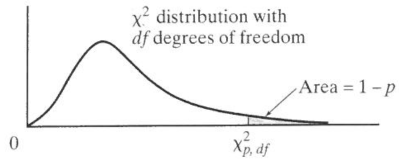
If H_0 is true then we know that d comes from a χ^2 distribution. So we can calculate the probability of finding a value as extreme as d, and use this to decide whether or not to reject the null hypothesis. The result is: we reject H_0 at significance level α if

$$d > \chi^2_{1-\alpha,m-1}$$

where the critical value $\chi^2_{1-\alpha,m-1}$ can be found from the chi-square tables. For example if $\alpha = 0.05$ (typical value) and $m = 10$ the value is

$$\chi^2_{0.95,9} = 16.919$$

Table A.3 Upper and Lower Percentiles of χ^2 Distributions



df	p = 0.010	0.025	0.050	0.10	0.90	0.95	0.975	0.99
1	0.000157	0.000982	0.00393	0.0158	2.706	3.841	5.024	6.635
2	0.0201	0.0506	0.103	0.211	4.605	5.991	7.378	9.210
3	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345
4	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277
5	0.554	0.831	1.145	1.610	9.236	11.070	12.832	15.086
6	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812
7	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475
8	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090
9	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666
10	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209
11	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725
12	3.571	4.404	5.226	6.304	18.549	21.026	23.336	26.217
13	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688
14	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141
15	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578
16	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000
17	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409
18	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805
19	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191
20	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566
21	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932
22	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.289
23	10.196	11.688	13.091	14.848	32.007	35.172	38.076	41.638
24	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980
25	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314
26	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642
27	12.879	14.573	16.151	18.114	36.741	40.113	43.194	46.963
28	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278
29	14.256	16.047	17.708	19.768	39.087	42.557	45.722	49.588
30	14.953	16.791	18.493	20.599	40.256	43.773	46.979	50.892
31	15.655	17.539	19.281	21.434	41.422	44.985	48.232	52.191
32	16.362	18.291	20.072	22.271	42.585	46.194	49.480	53.486
33	17.073	19.047	20.867	23.110	43.745	47.400	50.725	54.776
34	17.789	19.806	21.664	23.952	44.903	48.602	51.966	56.061

Example 4 *A die is rolled 30 times and the frequency of each outcome is recorded in the following table. Test at significance level* $\alpha = 0.05$ *to decide if the die is fair. So the null hypothesis* H_0 *is that each outcome has probability* $1/6$.

Outcome	*1*	*2*	*3*	*4*	*5*	*6*
Probability under H_0	1/6	1/6	1/6	1/6	1/6	1/6
Observed frequency	4	11	5	3	5	2
Expected frequency under H_0	5	5	5	5	5	5

The test statistic is

$$d = \sum_{i=1}^{6} \frac{(\text{Observed}_i - \text{Expected}_i)^2}{\text{Expected}_i} = 10$$

We have $\alpha = 0.05$ *and* $df = m - 1 = 5$ *so the critical value is*

$$\chi^2_{1-\alpha,df} = \chi^2_{0.95,5} = 11.07$$

Since $d < \chi^2_{1-\alpha,df}$ *we do* not *reject* H_0*, and so we accept that the die may be fair; that is we do* not *reject the null hypothesis.*