## Introduction to Statistical Ideas and Methods

## The $t$ Distribution

If our data are $n$ independent observations from a Normal distribution with mean $\mu$, then $t=\frac{\bar{x}-\mu}{s / \sqrt{n}}$ has a $t$ distribution with $n-1$ degrees of freedom.

Properties of the $t$ distribution:

- Symmetric
- Bell-shaped
- Centred at 0
- Has heavier tails than the Normal distribution, that is, the probability of being out in the tails is greater
- The smaller the degrees of freedom, the heavier the tails
- As the degrees of freedom gets large, the distribution gets closer to the standard Normal distribution

Figure 1 below shows probability density functions for the standard Normal distribution, and $t$ distributions with 1,5 , and 25 degrees of freedom. From the plot we can see that, as the degrees of freedom grows, the $t$ probability density functions get closer to the probability density function of the standard Normal distribution. The $t$ distribution with 25 degrees of freedom closely follows that of the Normal distribution. The heaver tails for smaller degrees of freedom can be seen by considering the area under the curves, corresponding to probabilities of intervals. For example, Figure 2 shows the probability of being greater than 2 for random variables with the standard Normal distribution, $t$ distribution with 1 degree of freedom, and $t$ distribution with 5 degrees of freedom, respectively. The probability of being greater than 2 decreases with the degrees of freedom.

To calculate probabilities and quantiles (values with a given probability of being less than it) from a $t$ distribution, we need either a table or statistical software. The last two pages of this document gives an example of a $t$-table. (This table is from OpenIntro Statistics https://www.openintro.org/stat/textbook.php?stat_book=os.) Some things to note about using this table:

- Since the table needs to include various degrees of freedom (df), the table only gives a few probability values.
- The probability is in the top margin of the table and the quantiles are in the body of the table.
- The table only gives the positive quantiles. For negative values, use the fact that $t$ distributions are symmetric about 0 .


Figure 1: Probability density functions for the standard Normal distribution, and $t$ distributions with 1,5 , and 25 degrees of freedom (df).


Figure 2: $P(X>2), P(Y>2)$ and $P(Z>2)$ where $X$ has the $t$ distribution with 1 degree of freedom, $Y$ has the $t$ distribution with 5 degrees of freedom, and $Z$ has the standard Normal distribution.

- As the degrees of freedom get larger, the quantiles decrease for the same tail probability. The final row of the table, with $\mathrm{df}=\infty$, gives quantiles from the standard Normal distribution.
- If you need to use a $t$ table and the degrees of freedom that you need is not in the table, use the closest degrees of freedom that is smaller than the degrees of freedom that you need. This is considered more conservative than rounding up. So, for example, if you needed values from a $t$ distribution with 66 degrees of freedom, you would use the row with 60 degrees of freedom.
- Here are some examples of reading values from this table. Suppose $X$ is a random variable with a $t$ distribution with 4 degrees of freedom.

1. What is $P(X<-2.13)$ ?

Answer: From the row in the $t$ table for 4 degrees of freedom, the one tail probability associated with the quantile 2.13 is 0.050 . So $P(X>2.13)=P(X<$ $-2.13)=0.050$.
2. What is $P(X<-2.00)$ ?

Answer: 2.00 is not a quantile in the table, but it is between 1.53 and 2.13 with associated one tail probabilities of 0.100 and 0.050 , respectively. So $P(X<-2.00)$ must be between these probabilities, that is $0.050<P(X<-2.00)<0.100$.
3. What is the value of $a$ such that $P(X>a)=0.010$ ? The one tail probability 0.010 has quantile 3.75. Since this a right-tail probability (greater than), we need the positive quantile. So $a=3.75$.

- The two tails probabilities in the table are useful for finding quantiles for confidence intervals. For example, if we had 9 degrees of freedom, and wanted to construct a $95 \%$ confidence interval, we need the quantile so that the area in the two tails is $1-0.95=0.05$. This would be 2.26 .


## t-Probability Table

$\underbrace{}_{-3}$


| one tail two tails | 0.100 | 0.050 | 0.025 | 0.010 | 0.005 |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | 0.200 | 0.100 | 0.050 | 0.020 | 0.010 |
| df | 3.08 | 6.31 | 12.71 | 31.82 | 63.66 |
|  | 1.89 | 2.92 | 4.30 | 6.96 | 9.92 |
|  | 1.64 | 2.35 | 3.18 | 4.54 | 5.84 |
|  | 1.53 | 2.13 | 2.78 | 3.75 | 4.60 |
|  | 1.48 | 2.02 | 2.57 | 3.36 | 4.03 |
| 6 | 1.44 | 1.94 | 2.45 | 3.14 | 3.71 |
| 7 | 1.41 | 1.89 | 2.36 | 3.00 | 3.50 |
| 8 | 1.40 | 1.86 | 2.31 | 2.90 | 3.36 |
| 9 | 1.38 | 1.83 | 2.26 | 2.82 | 3.25 |
| 10 | 1.37 | 1.81 | 2.23 | 2.76 | 3.17 |
| 11 | 1.36 | 1.80 | 2.20 | 2.72 | 3.11 |
| 12 | 1.36 | 1.78 | 2.18 | 2.68 | 3.05 |
| 13 | 1.35 | 1.77 | 2.16 | 2.65 | 3.01 |
| 14 | 1.35 | 1.76 | 2.14 | 2.62 | 2.98 |
| 15 | 1.34 | 1.75 | 2.13 | 2.60 | 2.95 |
| 16 | 1.34 | 1.75 | 2.12 | 2.58 | 2.92 |
| 17 | 1.33 | 1.74 | 2.11 | 2.57 | 2.90 |
| 18 | 1.33 | 1.73 | 2.10 | 2.55 | 2.88 |
| 19 | 1.33 | 1.73 | 2.09 | 2.54 | 2.86 |
| 20 | 1.33 | 1.72 | 2.09 | 2.53 | 2.85 |
| 21 | 1.32 | 1.72 | 2.08 | 2.52 | 2.83 |
| 22 | 1.32 | 1.72 | 2.07 | 2.51 | 2.82 |
| 23 | 1.32 | 1.71 | 2.07 | 2.50 | 2.81 |
| 24 | 1.32 | 1.71 | 2.06 | 2.49 | 2.80 |
| 25 | 1.32 | 1.71 | 2.06 | 2.49 | 2.79 |
| 26 | 1.31 | 1.71 | 2.06 | 2.48 | 2.78 |
| 27 | 1.31 | 1.70 | 2.05 | 2.47 | 2.77 |
| 28 | 1.31 | 1.70 | 2.05 | 2.47 | 2.76 |
| 29 | 1.31 | 1.70 | 2.05 | 2.46 | 2.76 |
| 30 | 1.31 | 1.70 | 2.04 | 2.46 | 2.75 |


| one tail <br> two tails | 0.100 | 0.200 | 0.100 | 0.050 | 0.020 |
| ---: | ---: | ---: | ---: | ---: | ---: |
| df | 31 | 1.31 | 1.70 | 2.04 | 2.45 |
| 32 | 1.31 | 1.69 | 2.04 | 2.45 | 2.74 |
| 33 | 1.31 | 1.69 | 2.03 | 2.44 | 2.73 |
| 34 | 1.31 | 1.69 | 2.03 | 2.44 | 2.73 |
| 35 | 1.31 | 1.69 | 2.03 | 2.44 | 2.72 |
| 36 | 1.31 | 1.69 | 2.03 | 2.43 | 2.72 |
| 37 | 1.30 | 1.69 | 2.03 | 2.43 | 2.72 |
| 38 | 1.30 | 1.69 | 2.02 | 2.43 | 2.71 |
| 39 | 1.30 | 1.68 | 2.02 | 2.43 | 2.71 |
| 40 | 1.30 | 1.68 | 2.02 | 2.42 | 2.70 |
| 41 | 1.30 | 1.68 | 2.02 | 2.42 | 2.70 |
| 42 | 1.30 | 1.68 | 2.02 | 2.42 | 2.70 |
| 43 | 1.30 | 1.68 | 2.02 | 2.42 | 2.70 |
| 44 | 1.30 | 1.68 | 2.02 | 2.41 | 2.69 |
| 45 | 1.30 | 1.68 | 2.01 | 2.41 | 2.69 |
| 46 | 1.30 | 1.68 | 2.01 | 2.41 | 2.69 |
| 47 | 1.30 | 1.68 | 2.01 | 2.41 | 2.68 |
| 48 | 1.30 | 1.68 | 2.01 | 2.41 | 2.68 |
| 49 | 1.30 | 1.68 | 2.01 | 2.40 | 2.68 |
| 50 | 1.30 | 1.68 | 2.01 | 2.40 | 2.68 |
| 60 | 1.30 | 1.67 | 2.00 | 2.39 | 2.66 |
| 70 | 1.29 | 1.67 | 1.99 | 2.38 | 2.65 |
| 80 | 1.29 | 1.66 | 1.99 | 2.37 | 2.64 |
| 90 | 1.29 | 1.66 | 1.99 | 2.37 | 2.63 |
| 100 | 1.29 | 1.66 | 1.98 | 2.36 | 2.63 |
| 150 | 1.29 | 1.66 | 1.98 | 2.35 | 2.61 |
| 200 | 1.29 | 1.65 | 1.97 | 2.35 | 2.60 |
| 300 | 1.28 | 1.65 | 1.97 | 2.34 | 2.59 |
| 400 | 1.28 | 1.65 | 1.97 | 2.34 | 2.59 |
| 500 | 1.28 | 1.65 | 1.96 | 2.33 | 2.59 |
| $\infty$ | 1.28 | 1.65 | 1.96 | 2.33 | 2.58 |
|  |  |  |  |  |  |

This table is from OpenIntro Statistics https://www.openintro.org/stat/textbook.php? stat_book=os.

