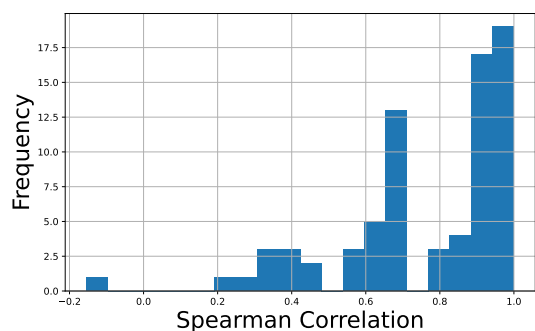
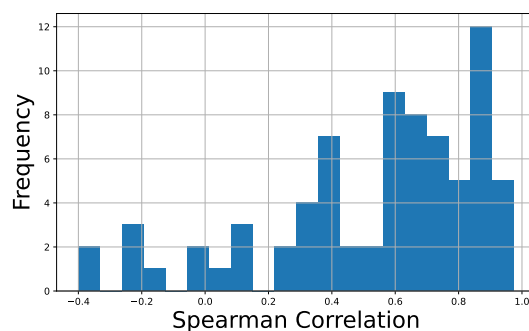


K Spearman Rank Distributions



(a) Humor. Model: Llama



(b) Humor | Persuasiveness. Model: Llama

Figure 13: Example of the distribution of Spearman correlation values between the desired levels and rank of the target concept. Result shown for argument generation task, using Llama 3-11B, with the target concept only and secondary concept level *randomly* sampled. For example, “Humor | Persuasiveness” denotes responses generated independently for each humor level (target concept) while persuasiveness is randomly set for each inference.